By preliminary amendment, Applicants correct certain erroneous references to the sequence identifiers and add sequence identifiers where appropriate throughout the specification. These corrections and additions are shown in detail in the marked-up version of the affected paragraphs attached hereto as "Appendix - 1. "Revisions Made to Specification and Claims." Claims 7-21 and 34-37 have been canceled, and claims 3-6 and 24, 29, 30 and 32 have been amended. The changes to the amended claims are also shown in detail in Appendix - 1. A complete set of the pending claims, as presently amended, is attached as "Appendix - 2. Complete Set of Pending Claims, as Amended on September 17, 2001."

In the "Notice to File Missing Parts of Nonprovisional Application," the Initial Patent Examiner stated that the application is informal for the reason that excessive text is contained in the drawings as originally filed. In reply, substitute sheets of the drawings are submitted containing Figures 1-12 to replace the set of drawings filed on April 10, 2001. As required by the Initial Patent Examiner, the Figures have been revised to omit certain text so as to better comply with 37 CFR § 1.84. Text deleted from the Figures has been moved to the corresponding figure descriptions in the "Brief Description of the Drawings" section of the Specification, or has been moved to the "Detailed Description" section where the particular Figure is discussed, or has been deleted because it is considered duplicative of description in the Specification as originally filed. A marked-up copy of the Figures is attached as "Appendix - 3. Details of Drawing Revisions Submitted on September 17, 2001," showing the proposed changes by circling the new text in red ink and by striking out deleted text in red ink. If the consent of the Examiner is necessary for entry of these substitute sheets of drawings, Applicants respectfully request approval by the Examiner of the proposed corrections to the Figures and entry of the substitute sheets of drawings.

Also enclosed is a computer diskette containing the initial computer readable form (CRF) of the Sequence Listing and a printed copy of the same, along with a statement that the content of the CRF and the original printed copy attached to the Application as filed is the same, and that the CRF and the attached printed copy are the same, and that no new matter is introduced thereby.

Applicants respectfully submit that the amendments to the specification, claims and drawings presented herein are to correct typographical errors, to facilitate examination of the application by ensuring compliance with 37 CFR § 1.821-1.824 or 37 CFR § 1.75(c) or 37 CFR § 1.84 prior to the first examination on the merits, or to comply with specific requirements stated in the "Notice to File Missing Parts." No new matter is introduced by way of these amendments.

If any additional fees are required or if the fees submitted are in excess of that required, please charge or credit the appropriate fees to Conley, Rose & Tayon, P.C. Deposit Account No.

03-2769/1889/00401/CGM. In the event that an extension of time is necessary in order for this submission to be deemed timely filed, please consider this a petition therefor and charge the associated fee to Conley, Rose & Tayon, P.C. Deposit Account No. 03-2769/1889/00401/CGM.

Respectfully submitted,

Carol G. Mintz

Carol G. Mintz
PTO Reg. No. 38,561
CONLEY, ROSE & TAYON
P.O. Box 3267
Houston, TX 77253-3267
(713) 238-8000 (Phone)
(713) 238-8008 (Fax)

APPENDIX - 1

REVISIONS MADE TO REPLACEMENT PARAGRAPHS AND CLAIMS

[0005] The invention relates to a compound that competitively inhibits binding of CSP **[SEQ ID NO:4]** to *S. mutans* histidine kinase **[SEQ ID NO:6]**. The compound is preferably a peptide or an antibody. The compound is preferably a derivative of ~~[SEQ ID NO:2]~~ **[SEQ ID NO:3]**, a fragment of ~~[SEQ ID NO:2]~~ **[SEQ ID NO:3]** or a derivative of a fragment of ~~[SEQ ID NO:2]~~ **[SEQ ID NO:3]**.

[0006] Schematic layout of the arrangement of the genetic locus encoding the signal peptide precursor (ComC) **[SEQ ID NO:1]**, the histidine kinase (ComD) **[SEQ ID NO:5]** and the response regulator (ComE) **[SEQ ID NO:7]**. Note that this arrangement is different from other loci in related streptococci for the following reasons: a) The comC gene **[SEQ ID NO:1]** is transcribed from its own unique promoter, unlike the genes thus far described in other streptococci that are arranged in an operon-like cluster with the comC/DE genes being transcribed from a single promoter.

[0007] b) The comC gene **[SEQ ID NO:1]** is separated by 148 nucleotides from the comD gene **[SEQ ID NO:5]**.

[0008] Shows the nucleic acid molecule that is ~~[SEQ ID NO 1]~~ **[SEQ ID NO:3]**. In a preferred embodiment, the figure shows CSP (competence signal peptide **[SEQ ID NO 3]**). Nucleotide sequence of the locus. Figure 2 also shows histidine kinase **[SEQ ID NO 5]** sequences and response regulator **[SEQ ID NO 7]** sequences.

[0008.5] Figure 2A. *S. mutans* comC gene [SEQ ID NO:1]. Encodes a precursor to a signal peptide [SEQ ID NO:2].

[0008.6] Figure 2B. *S. mutans* CSP encoding sequence [SEQ ID NO:3]. Encodes a Competence Signal Peptide [SEQ ID NO:4].

[0008.7] Figure 2C. *S. mutans* comD gene [SEQ ID NO:5]. Encodes a protein that functions as a histidine kinase receptor [SEQ ID NO:6].

[0008.8] Figure 2D. *S. mutans* comE gene [SEQ ID NO:7]. Encodes a response regulator that activates transcription of a number of genes [SEQ ID NO:8].

[[0009] Sequence of the deduced amino acid sequence of the signal peptide **[SEQ ID NO 4]**, histidine kinase **[SEQ ID NO: 6]**, and response regulator **[SEQ ID NO: 8]**.

[0009.4] Figure 3A. *S. mutans* ComC protein (CSP Precursor) [SEQ ID NO:2].

[0009.5] Figure 3B. *S. mutans* ComD protein (Histidine Kinase) [SEQ ID NO:6].

[0009.6] Figure 3C. *S. mutans* ComE protein (Response Regulator) [SEQ ID NO:8].

[0010] The deduced amino acid sequence of the signal peptide precursor **in various strains** and its predicted cleavage site~~, following the glycine-glycine signal~~ **The original peptide is expressed as a 46 amino acid peptide that is cleaved after the glycine-glycine residues to generate an active signal peptide**.

[0011] Shows the peptide that is ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]**. The synthetic signal peptide **[SEQ ID NO:16]** that is effective at inducing competence, biofilm formation and acid tolerance in *Streptococcus mutans*.

[0013] Table illustrating the effect of synthetic peptide on genetic competence in *S. mutans* cells. **Induction of genetic transformation in *Streptococcus mutans* by synthetic competence stimulating peptide (SCSP).**

[0015] ComCDE local region **[SEQ ID NO:23]**. **The ComC (first highlighted region; nucleotide 101 to 241), ComD (second highlighted region; nucleotides 383 to 1708) and ComE (third highlighted region; nucleotides 1705 to 2457) proteins are highlighted.**

[0016] The comX DNA sequence **[SEQ ID NO:22]**, protein sequence **[SEQ ID NO:23]**, and the comX gene local region **[SEQ ID NO:24]** with 100bp included both upstream and downstream (promoter is upstream).

[0016.4] Figure 10A. *S. mutans* comX gene [SEQ ID NO:22].

[0016.5] Figure 10B. *S. mutans* ComX protein [SEQ ID NO:25].

[0016.6] Figure 10C. *S. mutans* comX gene local region [SEQ ID NO:26].

[0017] The comA and comB nucleotide **[SEQ ID NO:25] and [SEQ ID NO:27]** and amino acid sequences **[SEQ ID NO:26] and [SEQ ID NO:28]** . ComA and ComB are the components of the CSP exporter.

[0017.5] Figure 11A. *S. mutans* comA gene [SEQ ID NO:27].

[0017.6] Figure 11B. *S. mutans* ComA protein [SEQ ID NO:28].

[0017.7] Figure 11C. *S. mutans* comB gene [SEQ ID NO:29].

[0017.8] Figure 11D. *S. mutans* ComB protein [SEQ ID NO:30].

[0018] Illustrates the effect of synthetic peptide on acid resistance tolerance in *S. mutans comC* deficient cells. **Addition of synthetic signal peptide (CSP) [SEQ ID NO:16] into the culture of the comC mutant restored the ability of the mutant to survive a low pH challenge when compared to the parent strain NG8.**

[0020] We have identified a genetic locus in *S. mutans* consisting of three genes that encode: 1) a peptide precursor **[SEQ ID NO:2]** that is processed during export into a secreted 21-amino acid peptide (CSP) **[SEQ ID NO:4]**; 2) a histidine kinase **[SEQ ID NO:6]** that acts as a cell surface receptor activated by the peptide; 3) a response regulator **[SEQ ID NO:8]** that activates a number of other genes involved in genetic competence, biofilm formation, and acid tolerance of *S. mutans*. These properties have been attributed to the bacterium's ability to cause dental caries. Inactivation of any of these three genes or impairment of interaction or activity of any of their encoded proteins will disrupt the bacterium's ability to take up foreign DNA, form biofilms, and tolerate acidic pH.

[0021] *Streptococcus mutans* is a resident of the biofilm environment of dental plaque, a matrix of bacteria and extracellular material that adheres to the tooth surface. Under appropriate environmental conditions populations of *S. mutans* and the pH of the surrounding plaque will drop. *S. mutans*, being among the most acid tolerant organisms residing in dental plaque, will increase it numbers in this acidic environment and eventually become a dominant member of the plaque community. This situation eventually leads to dissolution of the tooth enamel, resulting in the development of dental caries. We control the accumulation and acid tolerance of this bacterium to make it less able to cause caries. We accomplish this by using inhibitors of an extracellular signal peptide that promotes the expression of genes involved in *S. mutans* biofilm formation and acid tolerance. The invention includes compounds that inhibit the action of the peptide. These inhibitors can include peptides, antibodies, or other agents that specifically inhibit the activation of the histidine kinase and the family of genes activated as a result of the histidine kinase activation by the signal molecule. Inhibitors include: modified structures of the peptide where amino acids are removed from the N- and/or COOH terminal of the peptide and/or substitutions of internal amino acid residues. We delete, one, two to 5, 6 to 10 and 10 to 15 amino acids from the peptide (for example at either terminal) and measure competitive inhibition of signal peptide binding to histidine

[0022] kinase (1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more amino acids are deleted and inhibition measured). Inhibitors also include antibodies raised against the 21-amino acid CSP **[SEQ ID NO:4]** alone or coupled to a larger molecule to increase immunogenicity.

[0024] In addition to identifying the genes encoding this signaling/sensing system, we have identified and chemically synthesized a 21-amino acid peptide **[SEQ ID NO:16]** that promotes biofilm formation and acid tolerance of *S. mutans*. A survey of the literature and genome databases reveals that genes similar to this signal-receptor system are present in most Gram-

positive bacteria, and therefore an inhibitor, or family of related inhibitors may be effective at inhibiting biofilm formation among a large group of bacteria.

[0025] The invention treats or prevents dental caries by addition of compounds that inhibit the stimulatory action of the 21-amino acid peptide **[SEQ ID NO:4]** on biofilm formation and acid tolerance of *S. mutans*. This is accomplished by delivery of these compounds to the biofilm and/or to incorporate these inhibitors into materials to control growth on surfaces. This includes delivery by topical application, alone or in combination with other compounds including toothpaste, mouthwash, food or food additives.

[0027] The invention includes an isolated CSP from *S. mutans*. The invention also includes a recombinant isolated CSP **[SEQ ID NO:4]** peptide produced by a cell including a nucleic acid molecule encoding CSP **[SEQ ID NO:3]** operably linked to a promoter. The invention also includes an isolated nucleic acid molecule encoding a CSP **[SEQ ID NO:3]**. The peptide we work with is preferably chemically synthesized **[SEQ ID NO:16]**.

[0028] The invention includes CSP-encoding nucleic acid molecules **[SEQ ID NO:3]** and molecules having sequence identity or which hybridize to the CSP-encoding sequence and which encode a peptide having CSP activity (preferred percentages for sequence identity are described below) as well as vectors including these molecules. The invention also includes CSP **[SEQ ID NO:4]** or peptides having sequence identity (preferred percentages described below) or which have CSP activity. The nucleic acid molecules and peptides of the invention may be from S. mutans and they may be isolated from a native source, synthetic or recombinant. The invention includes CSP **[SEQ ID NO:4]** or peptides having sequence identity, which have CSP activity, as prepared by the processes described in this application.

[0029] The invention includes an isolated HK **[SEQ ID NO:6]** from *S. mutans.* The invention also includes a recombinant isolated HK polypeptide produced by a cell including a nucleic acid molecule encoding HK **[SEQ ID NO:5]** operably linked to a promoter. The invention also includes an isolated nucleic acid molecule encoding a HK polypeptide **[SEQ ID NO:6]**.

[0030] The invention includes HK-encoding nucleic acid molecules and molecules having sequence identity or which hybridize to the HK-encoding sequence **[SEQ ID NO:5]** and which encode a protein having HK activity (preferred percentages for sequence identity are described below) as well as vectors including these molecules. The invention also includes HK **[SEQ ID NO:4]** or polypeptides having sequence identity (preferred percentages described below) or which have HK activity. The nucleic acid molecules and polypeptides of the invention may be

from *S. mutans* and they may be isolated from a native source, synthetic or recombinant. The invention includes HK **[SEQ ID NO:4]** or polypeptides having sequence identity, which have HK activity, as prepared by the processes described in this application.

[0031] The invention includes an isolated RR **[SEQ ID NO:6]** from *S. mutans*. The invention also includes a recombinant isolated RR **[SEQ ID NO:6]** polypeptide produced by a cell including a nucleic acid molecule encoding RR **[SEQ ID NO:6]** operably linked to a promoter. The invention also includes an isolated nucleic acid molecule encoding a RR polypeptide.

[0032] The invention includes RR-encoding nucleic acid molecules and molecules having sequence identity or which hybridize to the RR-encoding sequence **[SEQ ID NO:5]** and which encode a polypeptide having RR activity (preferred percentages for sequence identity are described below) as well as vectors including these molecules. The invention also includes RR **[SEQ ID NO:6]** or polypeptides having sequence identity (preferred percentages described below) or which have RR activity. The nucleic acid molecules and polypeptides of the invention may be from *S. mutans* and they may be isolated from a native source, synthetic or recombinant. The invention includes RR **[SEQ ID NO:6]** or polypeptides having sequence identity, which have RR activity, as prepared by the processes described in this application.

[0033] The comA and comB nucleotide **[SEQ ID NO:27 and SEQ ID NO:29]** and amino acid sequences **[SEQ ID NO:28 and SEQ ID NO:30]** are also aspects of the invention. ComA and ComB are components of the CSP exporter. The discussion of variants, sequence identity etc. for CSP, HK, RR applies to both the full sequences shown in the figures as well as bracketed portions of sequences (coding regions). The peptides and polypeptides may be natural, recombinantly produced or synthetic.

[0034] The invention includes nucleic acid molecules that are functional equivalents of all or part of the CSP sequence in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. (A nucleic acid molecule may also be referred to as a DNA sequence or nucleotide sequence in this application. All these terms have the same meaning as nucleic acid molecule). Functionally equivalent nucleic acid molecules are DNA and RNA (such as genomic DNA, complementary DNA, synthetic DNA, and messenger RNA molecules) that encode peptides having the same or similar CSP activity as the CSP peptide shown in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]**. Functionally equivalent nucleic acid molecules can encode peptides that contain a region having sequence identity to a region of a CSP peptide **[SEQ ID NO:4]** or more preferably to the entire CSP peptide. Identity is calculated according to methods known in the art. The ClustalW program (preferably using default parameters) [Thompson, JD et al., Nucleic Acid Res. 22:4673-4680.], described below, is most preferred. For example, if a nucleic acid molecule (called "Sequence A") has 90% identity to a

portion of the nucleic acid molecule in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**, then Sequence A will preferably be identical to the referenced portion of the nucleic acid molecule in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**, except that Sequence A may include up to 10 point mutations, such as substitutions with other nucleotides, per each 100 nucleotides of the referenced portion of the nucleic acid molecule in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. Mutations described in this application preferably do not disrupt the reading frame of the coding sequence. Nucleic acid molecules functionally equivalent to the CSP sequences can occur in a variety of forms as described below.

[0035] Nucleic acid molecules may encode conservative amino acid changes in CSP peptide **[SEQ ID NO:4]**. The invention includes functionally equivalent nucleic acid molecules that encode conservative amino acid changes within a CSP amino acid sequence and produce silent amino acid changes in CSP.

[0036] Nucleic acid molecules may encode non-conservative amino acid substitutions, additions or deletions in CSP peptide. The invention includes functionally equivalent nucleic acid molecules that make non-conservative amino acid changes within the CSP amino acid sequence in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]**. Functionally equivalent nucleic acid molecules include DNA and RNA that encode peptides, peptides and proteins having non-conservative amino acid substitutions (preferably substitution of a chemically similar amino acid), additions, or deletions but which also retain the same or similar CSP activity as the CSP peptide shown in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]**. The DNA or RNA can encode fragments or variants of CSP. Fragments are useful as immunogens and in immunogenic compositions (U.S. Patent No. 5,837,472). The CSP or CSP-like activity of such fragments and variants is identified by assays as described below. Fragments and variants of CSP encompassed by the present invention should preferably have at least about 40%, 60%, 80% or 95% sequence identity to the naturally occurring CSP nucleic acid molecule, or a region of the sequence, such as the coding sequence or one of the conserved domains of the nucleic acid molecule, without being identical to the sequence in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. Sequence identity is preferably measured with the ClustalW program (preferably using default parameters) (Thompson, JD et al., Nucleic Acid Res. 22:4673-4680)

[0037] Nucleic acid molecules functionally equivalent to the CSP nucleic acid molecule in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** will be apparent from the following description. For example, the sequence shown in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** may have its length altered by natural or artificial mutations such as partial nucleotide insertion or deletion, so that when the entire length of the coding sequence within ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**, is taken as 100%, the functional

equivalent nucleic acid molecule preferably has a length of about 60-120% thereof, more preferably about 80-110% thereof. Fragments may be less than 60%.

[0038] Nucleic acid molecules containing partial (usually 80% or less, preferably 60% or less, more preferably 40% or less of the entire length) natural or artificial mutations so that some codons in these sequences code for different amino acids, but wherein the resulting peptide retains the same or similar CSP activity as that of a naturally occurring CSP peptide **[SEQ ID NO:4]**. The mutated DNAs created in this manner should preferably encode a peptide having at least about 40%, preferably at least about 60%, at least about 80%, and more preferably at least about 90% or 95% sequence identity to the amino acid sequence of the CSP peptide in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]**. The ClustalW program preferably assesses sequence identity.

[0039] Since the genetic code is degenerate, the nucleic acid sequence in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** is not the only sequence which may code for a peptide having CSP activity. This invention includes nucleic acid molecules that have the same essential genetic information as the nucleic acid molecule described in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. Nucleic acid molecules (including RNA) having one or more nucleic acid changes compared to the sequences described in this application and which result in production of a peptide shown in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:4]** are within the scope of the invention.

[0040] Other functional equivalent forms of CSP-encoding nucleic acids can be isolated using conventional DNA-DNA or DNA-RNA hybridization techniques. Thus, the present invention also includes nucleic acid molecules that hybridize to one or more of the sequences in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** or its complementary sequence, and that encode expression for peptides, peptides and proteins exhibiting the same or similar activity as that of the CSP peptide produced by the DNA in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** or its variants. Such nucleic acid molecules preferably hybridize to the sequence in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** under moderate to high stringency conditions (see Sambrook et al. Molecular Cloning: A Laboratory Manual, Most Recent Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). High stringency washes have low salt (preferably about 0.2% SSC), and low stringency washes have high salt (preferably about 2% SSC). A temperature of about 37 °C or about 42 °C is considered low stringency, and a temperature of about 50-65 °C is high stringency. The invention also includes a method of identifying nucleic acid molecules encoding a CSP activator peptide (preferably a mammalian peptide), including contacting a sample containing nucleic acid molecules including all or part of ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** (preferably at least about 15 or 20 nucleotides of ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**) under moderate or high stringency hybridization conditions and identifying nucleic acid molecules which hybridize to the nucleic

acid molecules including all or part of ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**.). Similar methods are described in U.S. Patent No. 5,851,788, which is incorporated by reference in its entirety.

[0041] The invention also includes methods of using all or part of the nucleic acid molecules which hybridize to all or part of ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**, for example as probes or in assays to identify antagonists or inhibitors of the peptides produced by the nucleic acid molecules (described below). The invention also includes methods of using nucleic acid molecules having sequence identity to the CSP nucleic acid molecule **[SEQ ID NO:3]** (as described below) in similar methods.

[0042] The invention also includes a nucleic acid molecule detection kit including, preferably in a suitable container means or attached to a surface, a nucleic acid molecule of the invention encoding CSP **[SEQ ID NO:4]** or a peptide having CSP activity and a detection reagent (such as a detectable label). Other variants of kits will be apparent from this description and teachings in patents such as U.S. Patent Nos. 5,837,472 and 5,801,233,which are incorporated by reference in their entirety.

[0043] A nucleic acid molecule described above is considered to have a function substantially equivalent to the CSP nucleic acid molecules **[SEQ ID NO:3]** of the present invention if the peptide **[SEQ ID NO:4]** produced by the nucleic acid molecule has CSP activity. A peptide has CSP activity if it can stimulate genetic competence and acid tolerance in *S. mutans*. Activation of the HK **[SEQ ID NO:6]**/RR **[SEQ ID NO:8]** is shown where a peptide is capable of stimulating the uptake and incorporation of foreign DNA. We describe below how the activity of these peptide-mediated processes can be measured by determining the efficiency of plasmid uptake, which is a measure of genetic competence. Since the ability to transport and incorporate foreign DNA relies on activation of the HK **[SEQ ID NO:6]**/RR **[SEQ ID NO:8]** and subsequent genes activated by the signal cascade initiated by the signal peptide, measurement of the conferment of erythromycin resistance by cells exposed to the peptide and plasmid DNA conferring erythromycin resistance indicates its level of function. Conversely if an inhibitor is capable of interfering with the action of the peptide the competence assay will indicate this by a corresponding decrease in the number of cells that acquire erythromycin resistance as described in the assays below (assays of genetic competence and assay of transformation of biofilms). Activation of the HK **[SEQ ID NO:6]**/RR **[SEQ ID NO:8]** is also shown where a peptide is capable of stimulating an acid tolerance response. We describe below how the activity of these peptide-mediated processes can be measured by determining the survival rate of cells in acidic pH conditions. Since the ability to survive exposure to acidic pH depends on the activation of the HK/RR and subsequent genes activated by the signal peptide, measurement of the survival of *S. mutans* in low pH conditions indicates the level of function of

the signal peptide. Conversely, if an inhibitor is capable of interfering with the signal peptide sensing system the assay for acid adaptation will indicate this by a corresponding decrease in the survival rate of cells grown in acidic pH conditions as described in the assay below (assay of acid adaptation).

[0049] Another embodiment of the invention relates to a method of transfecting a cell with a nucleic acid molecule of the invention, inserted in an expression vector to produce a cell expressing the CSP peptide **[SEQ ID NO:4]** or other peptide of the invention. The invention also relates to a method of expressing the peptides of the invention in a cell. A preferred process would include culturing a cell including a recombinant DNA vector including a nucleic acid molecule encoding CSP **[SEQ ID NO:3]** (or another nucleic acid molecule of the invention) in a culture medium so that the peptide is expressed. The process preferably further includes recovering the peptide from the cells or culture medium.

[0050] The invention also includes oligonucleotide probes made from the cloned CSP nucleic acid molecules described in this application or other nucleic acid molecules of the invention (see Materials and Methods section). The probes may be 15 to 20 nucleotides in length. A preferred probe is at least 15 nucleotides of CSP in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. The invention also includes at least 15 consecutive nucleotides of ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]**. The probes are useful to identify nucleic acids encoding CSP peptides as well as peptides functionally equivalent to CSP. The oligonucleotide probes are capable of hybridizing to the sequence shown in ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** under stringent hybridization conditions. A nucleic acid molecule encoding a peptide of the invention may be isolated from other organisms by screening a library under moderate to high stringency hybridization conditions with a labeled probe. The activity of the peptide encoded by the nucleic acid molecule is assessed by cloning and expression of the DNA. After the expression product is isolated, the peptide is assayed for CSP activity as described in this application.

[0051] Functionally equivalent CSP nucleic acid molecules from other cells, or equivalent CSP-encoding cDNAs or synthetic DNAs, can also be isolated by amplification using Polymerase Chain Reaction (PCR) methods. Oligonucleotide primers, such as degenerate primers, based on ~~[SEQ ID NO:1]~~ **[SEQ ID NO:3]** can be prepared and used with PCR and reverse transcriptase (E. S. Kawasaki (1990), In Innis et al., Eds., PCR Protocols, Academic Press, San Diego, Chapter 3, p. 21) to amplify functional equivalent DNAs from genomic or cDNA libraries of other organisms. The oligonucleotides can also be used as probes to screen cDNA libraries.

[0052] The present invention includes not only the peptides encoded by the sequences of the invention, but also functionally equivalent peptides, peptides and proteins that exhibit the same or similar CSP peptide activity. A peptide is considered to possess a function substantially equivalent to that of the CSP peptide **[SEQ ID NO:4]** if it has CSP activity. CSP activity means that it is able to confer genetic competence to *S. mutans*, as measured by an increased ability to incorporate and express foreign genetic material, when added to cells as described in the assay of genetic competence below. CSP activity also means that the peptide is able to confer an acid tolerance response in *S. mutans* as measured by an increase in cell survival under acidic pH conditions when added to cells as described in the assay for acid adaptation below. Functionally equivalent peptides, peptides and proteins include peptides, peptides and proteins that have the same or similar protein activity as CSP when assayed, i.e. they are able to stimulate genetic competence and low pH tolerance (the ability to withstand acid challenges of pH 3.5 –pH 3.0 for up to 3 hours) in *S. mutans*. A peptide has CSP activity if it is capable of increasing the frequency of uptake and expression of foreign DNA as described in the following assay for genetic competence and if the peptide can promote an acid tolerance response as described in the assay for acid adaptation.

[0053] Identity refers to the similarity of two peptides or proteins that are aligned so that the highest order match is obtained. Identity is calculated according to methods known in the art, such as the ClustalW program. For example, if a peptide (called "Sequence A") has 90% identity to a portion of the peptide in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:16]**, then Sequence A will be identical to the referenced portion of the peptide in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:16]**, except that Sequence A may include up to 1 point mutations, such as substitutions with other amino acids, per each 10 amino acids of the referenced portion of the peptide in ~~[SEQ ID NO:2]~~ **[SEQ ID NO:16]**. Peptides, peptides and proteins functional equivalent to the CSP peptides can occur in a variety of forms as described below.

[0054] Peptides biologically equivalent in function to CSP peptide include amino acid sequences containing amino acid changes in the CSP sequence **[SEQ ID NO:4]**. The functional equivalent peptides have at least about 40% sequence identity, preferably at least about 60%, at least about 75%, at least about 80%, at least about 90% or at least about 95% sequence identity, to the natural CSP peptide **[SEQ ID NO:4]** or a corresponding region. The ClustalW program preferably determines sequence identity. Most preferably, 1, 2, 3, 4, 5, 5-10, 10-15 amino acids are modified.

[0058] The variants preferably retain the same or similar CSP activity as the naturally occurring CSP **[SEQ ID NO:4]**. The CSP activity of such variants can be assayed by techniques described in this application and known in the art.

[0059] Variants produced by combinations of the techniques described above but which retain the same or similar CSP activity as naturally occurring CSP **[SEQ ID NO:4]** are also included in the invention (for example, combinations of amino acid additions, and substitutions).

[0062] As well, probes and antibodies for a histidine kinase **[SEQ ID NO:5 and SEQ ID NO:6]**, response regulator **[SEQ ID NO:7 and SEQ ID NO:8]** comA **[SEQ ID NO:27 and SEQ ID NO:28]** or comB **[SEQ ID NO:29 and SEQ ID NO:30]** may be prepared using the description in this application and techniques known in the art. The description for preparation of CSP variants and mutants is also applicable to the histidine kinase **[SEQ ID NO:5 and SEQ ID NO:6]**, response regulator **[SEQ ID NO:7 and SEQ ID NO:8]** or comA **[SEQ ID NO:27 and SEQ ID NO:28]** and comB **[SEQ ID NO:29 and SEQ ID NO:30]** of the invention. The invention also includes fragments of HK **[SEQ ID NO:5 and SEQ ID NO:6]** having HK activity, fragments of RR **[SEQ ID NO:7 and SEQ ID NO:8]** having RR activity and fragments of comA **[SEQ ID NO:27 and SEQ ID NO:28]** or comB **[SEQ ID NO:29 and SEQ ID NO:30]** having activity.

[0063] The activity of the CSP peptide **[SEQ ID NO:4]** may be varied by carrying out selective site-directed mutagenesis. We characterize the binding domain and other critical amino acid residues in the peptide that are candidates for mutation, insertion and/or deletion. Sequence variants may be synthesized. A DNA plasmid or expression vector containing the CSP nucleic acid molecule **[SEQ ID NO:3]** or a nucleic acid molecule having sequence identity may be used for these studies using the U.S.E. (Unique site elimination) mutagenesis kit from Pharmacia Biotech or other mutagenesis kits that are commercially available, or using PCR. Once the mutation is created and confirmed by DNA sequence analysis, the mutant peptide is expressed using an expression system and its activity is monitored. This approach is useful to identify CSP inhibitors. All these modifications of the CSP DNA sequences **[SEQ ID NO:3]** presented in this application and the peptides produced by the modified sequences are encompassed by the present invention.

[0064] The CSP inhibitors are also useful when combined with a carrier in a pharmaceutical composition. The compositions are useful when administered in methods of medical treatment or prophylaxis of a disease, disorder or abnormal physical state caused by *S. mutans*. The invention also includes methods of medical treatment of a disease, disorder or abnormal

physical state characterized by excessive *S. mutans* or levels or activity of CSP peptide **[SEQ ID NO:4]**, for example by administering a pharmaceutical composition including a carrier and a CSP inhibitor. Caries is one example of a disease, which can be treated or prevented by antagonizing CSP **[SEQ ID NO:4]**.

[0066] CSP activity could be blocked by antisense mRNA or by inhibiting the activity of the exporter that secretes it from the cell. We have the sequence of these exporters. There are two copies of the genes (comAB) **[SEQ ID NO:27 and SEQ ID NO:29]** that are involved in export.

[0067] Nucleic acid molecules (antisense inhibitors of CSP) **[SEQ ID NO:3]** and competitive inhibitors of CSP **[SEQ ID NO:4]** may be introduced into cells using *in vivo* delivery vehicles such as liposomes. They may also be introduced into these cells using physical techniques such as microinjection and electroporation or chemical methods such as coprecipitation or using liposomes.

[0070] Antibodies directed against the CSP **[SEQ ID NO:4 or SEQ ID NO:16]** would provide protection against caries. Antibodies may be manufactured as described below. Alternatively, a peptide of the invention **[SEQ ID NO:4 or SEQ ID NO:16]** or a fragment thereof may be used with a carrier to make a vaccine. The peptide or fragment may also be conjugated to another molecule to increase its antigenicity. Antibodies can also be coupled to the peptide (Brady, L.J. et al., "Monoclonal Antibody-Mediated Modulation of the Humoral Immune Response against Mucosally Applied *Streptococcus mutans*" (in press). In order to enhance the immune response the peptide can be coupled to KLH, ovalbumin, or thyroglobulin prior to immunization. The vaccine composition will trigger the mammal's immune system to produce antibodies. The invention includes vaccine compositions and methods of vaccinating a mammal, preferably a human, against dental caries by administering to the mammal an effective amount of a vaccine composition. Techniques for preparing and using vaccines are known in the art. To prepare the vaccine, the peptide, or a fragment of the peptide, may be mixed with other antigens (of different immunogenicity), a vehicle or an excipient. Examples of peptide vaccines are found in U.S. Patent Nos. 5,679,352, 5,194,254 and 4,950,480. Techniques for preparing vaccines involving site-directed mutagenesis are described in U.S. Patent Nos. 5,714,372, 5,543,302, 5,433,945, 5,358,868, 5,332,583, 5,244,657, 5,221,618, 5,147,643, 5,085,862 and 5,073,494. Vaccines may be administered by known techniques, such as topical or parenteral administration. Vast changes are taking place in vaccinology consequent to the introduction of new technologies. Acellular purified fractions devoid of side effects, non-pathogenic but immunogenic mutants, recombinant technology, conjugated vaccines, combination vaccines (to limit the number of injections). Vaccine delivery systems can deliver multiple doses of the vaccine at a single

contact point. A genetically engineered oral vaccine is useful to impart better and longer duration of immunity. Oral vaccines are useful. The nose as a route for immunization is also useful. DNA alone can constitute the vaccines, inducing both humoral and cell-mediated immune responses. Live recombinant vaccines are also useful. Potent adjuvants add to the efficacy of the vaccines. One can also 'humanize' mouse monoclonals by genetic engineering and express these efficiently in plants. These recombinant antibodies are opening out an era of highly specific and safe therapeutic interventions. An advantage of preformed antibodies directed at a defined target and given in adequate amounts is the certainty of efficacy in every recipient, in contrast to vaccines, where the quality and quantum of immune response varies from individual to individual. For example, nasal immunization may be done as described in C. Jespersgaard et al. "Protective Immunity against *Streptococcus mutans* Infection in Mice after Intranasal Immunization with the Glucan-Binding Region of *S. mutans* Glucosyltransferase" Infection and Immunity, December 1999, p. 6543-6549, Vol. 67, No. 12. Vaccine compositions may comprise solid or liquid formulations such as gels, sprays, inhalants, tablets, toothpastes, mouthwashes or chewing gum.

[0072] Inhibitors are preferably directed towards CSP **[SEQ ID NO:4 or SEQ ID NO:16]** to block *S. mutans* competence, low pH tolerance and biofilm formation.

[0073] A method of identifying a compound which reduces the interaction of CSP **[SEQ ID NO:4 or SEQ ID NO:16]** with HK **[SEQ ID NO:6]**, can include: contacting (i) CSP **[SEQ ID NO:4 or SEQ ID NO:16]** with (ii) HK **[SEQ ID NO:6]**, a CSP-binding fragment of HK **[SEQ ID NO:6]** or a derivative of either of the foregoing in the presence of the compound; and b) determining whether the interaction between (i) and (ii) is reduced, thereby indicating that the compound reduces the interaction of CSP **[SEQ ID NO:4 or SEQ ID NO:16]** and HK **[SEQ ID NO:6]**. A CSP inhibitor (caries treating or preventing compound) inhibits the interaction between (i) and (ii). By way of example, one can screen a synthetic peptide library. One could also screen small non-peptide organic molecules.

[0074] In one embodiment, the invention includes an assay for evaluating whether test compounds are capable of acting as agonists or antagonists for CSP **[SEQ ID NO:4]**, or a peptide having CSP functional activity, including culturing cells containing DNA which expresses CSP **[SEQ ID NO:3]**, or a peptide having CSP activity so that the culturing is carried out in the presence of at least one compound whose ability to modulate CSP activity is sought to be determined and thereafter monitoring the cells for either an increase or decrease in the level of CSP **[SEQ ID NO:4 or SEQ ID NO:16]** or CSP activity. Other assays (as well as variations of

the above assay) will be apparent from the description of this invention and techniques such as those disclosed in U.S. Patent No. 5,851,788, 5,736,337 and 5,767,075 which are incorporated by reference in their entirety. For example, the test compound levels may be either fixed or variable.

[0075] The CSP **[SEQ ID NO:4 or SEQ ID NO:16]** peptide is also useful as an antigen for the preparation of antibodies that can be used to purify or detect other CSP-like peptides. Antibodies may also block CSP **[SEQ ID NO:4]** binding to HK **[SEQ ID NO:6]**. Antibodies are preferably targeted to the entire CSP **[SEQ ID NO:4]** sequence. The CSP peptide **[SEQ ID NO:4 or SEQ ID NO:16]** may be conjugated to other compounds, in order to increase immunogenicity.

[0076] We generate polyclonal antibodies against the CSP **[SEQ ID NO:4 or SEQ ID NO:16]**, which is a unique sequence. Monoclonal and polyclonal antibodies are prepared according to the description in this application and techniques known in the art. For examples of methods of preparation and uses of monoclonal antibodies, see U.S. Patent Nos. 5,688,681, 5,688,657, 5,683,693, 5,667,781, 5,665,356, 5,591,628, 5,510,241, 5,503,987, 5,501,988, 5,500,345 and 5,496,705, which are incorporated by reference in their entirety. Examples of the preparation and uses of polyclonal antibodies are disclosed in U.S. Patent Nos. 5,512,282, 4,828,985, 5,225,331 and 5,124,147 which are incorporated by reference in their entirety. Antibodies recognizing CSP **[SEQ ID NO:4 or SEQ ID NO:16]** can be employed to screen organisms or tissues containing CSP peptide **[SEQ ID NO:4]** or CSP-like peptides. The antibodies are also valuable for immuno-purification of CSP **[SEQ ID NO:4]** or CSP-like peptides from crude extracts.

[0077] An antibody (preferably the antibody described above) may be used to detect CSP **[SEQ ID NO:4]** or a similar peptide, for example, by contacting a biological sample with the antibody under conditions allowing the formation of an immunological complex between the antibody and a peptide recognized by the antibody and detecting the presence or absence of the immunological complex whereby the presence of CSP **[SEQ ID NO:4]** or a similar peptide is detected in the sample. The invention also includes compositions preferably including the antibody, a medium suitable for the formation of an immunological complex between the antibody and a peptide recognized by the antibody and a reagent capable of detecting the immunolgical complex to ascertain the presence of CSP **[SEQ ID NO:4]** or a similar peptide. The invention also includes a kit for the *in vitro* detection of the presence or absence of CSP **[SEQ ID NO:4]** or a similar peptide in a biological sample, wherein the kit preferably includes an antibody, a medium suitable for the formation of an immunological complex between the antibody and a peptide recognized by the antibody and a reagent capable of detecting the

immunological complex to ascertain the presence of CSP **[SEQ ID NO:4]** or a similar peptide in a biological sample. Further background on the use of antibodies is provided, for example in U.S. Patent Nos. 5,695,931 and 5,837,472, which are incorporated by reference in their entirety.

[0078] The ability of the peptide to activate the HK **[SEQ ID NO:6]** and RR **[SEQ ID NO:8]** and the subsequent genes involved in the conferral of the properties of genetic competence, acid tolerance and biofilm formation can be determined by measuring the efficiency of uptake and expression of DNA (preferably plasmid DNA) in *S. mutans* when exposed to signal peptide and/or inhibitor. Two methods modified based on the protocols described by Perry *et al.* Infect Immun, 41:722-727 and Lindler and Macrina J Bacteriol, 166:658-665 are used to assay genetic competence. The method involves adding DNA and CSP **[SEQ ID NO:3]** (preferably plasmid DNA) to a S. mutans culture (or culture of a bacteria expressing CSP **[SEQ ID NO:3]** or a variant thereof). The rate of transformation is then determined. *S. mutans* is preferably grown in THYE plus 5% horse serum (THYE-HS). After 2-hr incubation, 1 μg/ml plasmid DNA or 10 μg/ml of chromosomal DNA is added to the culture. To assay induction of competence, competence signal peptide, (SCSP) **[SEQ ID NO:16]** is then added to the cultures, incubation continued for 30 minutes with a final concentration of 500 ng/ml of SCSP added to each sample. After the 30-minute incubation equal amounts of DNA is added to each well (1 μg/ml plasmid or 10 μg/ml of chromosomal DNA) and incubation continued for another 2 hrs.. Cell dilutions were immediately spread on THYE agar plates plus appropriate antibiotics. Transformation frequency was expressed as the number of transformants (antibiotic resistant cells) per number of viable recipients. This is determined by comparing the number of cells able to grow in the presence of antibiotic (conferred by the applied plasmid or chromosomal DNA) relative to the total number of cells present (i.e., that grow in the absence of antibiotic). A higher value indicates a higher rate of transformation and thus is reflective of a stimulatory effect by the peptide. Consequently, addition of a molecule that successfully acts as an inhibitor results in a lower ratio of transformants/recipients, indicating that the inhibitor is effective at blocking activity of the CSP **[SEQ ID NO:4]**. CSP deficient cells **[SEQ ID NO:3 or SEQ ID NO:4]** may also be used in a variation of these assays. One can identify compounds that inhibit CSP **[SEQ ID NO:4]** or variants thereof by adding a test compound to the mixture to determine if the rate of transformation is decreased by the addition of the test compound.

[0079] The activity of the system can also be measured by an *in vitro* assay that relies on the measurement of marker protein expression (such as green fluorescent protein (GFP)) via expression from a fusion to a promoter controlled by the signal cascade initiated by CSP **[SEQ ID NO:4]**/HK **[SEQ ID NO:6]**/RR **[SEQ ID NO:8]**. One such promoter occurs immediately 5' proximal to the *S. mutans* comX gene. *S. mutans* cells grown in microtiter wells are exposed to

the CSP **[SEQ ID NO:4]** and/or inhibitor and the level of fluorescence of the comX::GFP strain is measured to give a quantitative measure of CSP **[SEQ ID NO:4]** stimulation (and conversely inhibitor activity). One can identify compounds that inhibit CSP **[SEQ ID NO:4]** or variants thereof by adding a test compound to the mixture to determine if the quantitative measure of CSP **[SEQ ID NO:4]** stimulation is decreased by the addition of the test compound.

[0080] The ability of CSP **[SEQ ID NO:4]** to promote acid resistance tolerance is determined by measuring the cell survival rate of *S. mutans* when exposed to acidic pH. In one example, *S. mutans* are first grown in batch culture to assay acid tolerance response in 'standard' log- and stationary-phase cells by using a modification of methods described previously by Svensäter *et al.* Oral Microbiol. Immunol., 12:266-73. Mid-log-phase cells are obtained by transferring one volume of overnight culture into nine volumes (1:10) of fresh TYG medium (pH 7.5) and incubated at 37^{0}C with 5% CO_2 for 2 hours. These cells are then collected by centrifugation at 8,000 x g for 10 min and resuspended in 2 ml of fresh TYG (pH 5.5) at various cell densities as determined by O.D_{600}. The cells are induced for acid adaptation by incubation at pH 5.5 for 2 h at 37^{0}C with 5% CO_2. The adapted log-phase cells are then exposed to the killing pH. Killing pH is pre-determined by incubating unadapted, mid-log phase cells in TYG medium at pH values from 6.0 to 2.0. Stationary-phase cells are prepared by re-suspending late-log phase cells in TY medium (tryptone-yeast extract) without glucose. The culture is incubated at 37^{0}C for 2 h to allow the cells to fully enter into stationary phase. Induction of acid adaptation in stationary-phase cells follows a similar procedure to that for log-phase cells. Adaptation of both log- and stationary-phase cells to acidic pH is determined by measuring the ability of bacterial cells to survive a killing pH for 3 h. Acid killing is initiated by resuspending cells in the same volume of fresh TYG (pH 3.5) and an aliquot of cell suspension is taken immediately from each sample to determine total viable cell number at zero time. The cells are then incubated for 3 h at 37^{0}C with 5% CO_2 and an aliquot of sample is taken to determine survival rate by viable cell counts. Addition of a molecule that successfully acts as an inhibitor results in a decrease in the acid resistance tolerance of *S mutans* resulting in a corresponding decrease in cell survival indicating that the inhibitor is effective at blocking activity of CSP **[SEQ ID NO:4]**. CSP **[SEQ ID NO:3 or SEQ ID NO:4]** deficient cells may also be used in a variation of these assays wherein addition of the signal peptide can complement the acid-adaptation-defective phenotype of a comC **[SEQ ID NO:1 or SEQ ID NO:2]** deficient cell. One can identify compounds that inhibit CSP **[SEQ ID NO:3 or SEQ ID NO:4]** or variants thereof by adding a test compound to the mixture to determine if the survival rate of cells is decreased by the addition of the test compound

[0081] Cells transformed with a nucleic acid molecule of the invention (histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:7]**) are useful as research tools. For example, one may obtain a cell (or a cell line, such as an immortalized cell culture or a primary cell culture) that does not express histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:8]**, insert a histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:7]** nucleic acid molecule in the cell, and assess the level of expression and activity. Alternatively, histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:7]** nucleic acid molecules may be over-expressed in a cell that expresses a histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:7]** nucleic acid molecule. In another example, experimental groups of cells may be transformed with vectors containing different types of histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator **[SEQ ID NO:7]** nucleic acid molecules to assess the levels of polypeptides and peptides produced, its functionality and the phenotype of the cells. The polypeptides and peptides are also useful for *in vitro* analysis of histidine **[SEQ ID NO:6]** kinase, CSP **[SEQ ID NO:4]** or response regulator **[SEQ ID NO:8]** activity or structure. For example, the polypeptides and peptides produced can be used for microscopy or X-ray crystallography studies.

[0082] The histidine kinase **[SEQ ID NO:5 and SEQ ID NO:6]**, CSP **[SEQ ID NO:3 and SEQ ID NO:4]** or response regulator **[SEQ ID NO:7 and SEQ ID NO:8]** nucleic acid molecules and polypeptides are also useful in assays for the identification and development of compounds to inhibit and/or enhance polypeptide or peptide function directly. For example, they are useful in an assay for evaluating whether test compounds are capable of acting as antagonists for histidine kinase **[SEQ ID NO:6]**, CSP **[SEQ ID NO:4]** or response regulator **[SEQ ID NO:8]** by: (a) culturing cells containing a nucleic acid molecule which expresses histidine kinase **[SEQ ID NO:5]**, CSP **[SEQ ID NO:3]** or response regulator peptides **[SEQ ID NO:7]** (or fragments or variants thereof having histidine **[SEQ ID NO:6]** kinase, CSP or response regulator activity) wherein the culturing is carried out in the presence of increasing concentrations of at least one test compound whose ability to inhibit histidine **[SEQ ID NO:6]** kinase, CSP **[SEQ ID NO:4]** or response regulator **[SEQ ID NO:8]** is sought to be determined; and (b) monitoring in the cells the level of inhibition as a function of the concentration of the test compound, thereby indicating the ability of the test compound to inhibit histidine kinase **[SEQ ID NO:6]**, CSP **[SEQ ID NO:4]** or response regulator **[SEQ ID NO:8]** activity.

[0085] Biofilms are developed on polystyrene microtiter plates to provide a rapid and simple method for assaying biofilm formation, and hence activity of the peptide **[SEQ ID NO:4]**/receptor

[SEQ ID NO:8]/kinase [SEQ ID NO:6] system. Formation of biofilms is initiated by inoculating 20 ul of cell suspension into each well containing 2 ml of biofilm medium (4X diluted Todd-Hewitt Yeast Extract supplemented with final concentration of 0.01% hog gastric mucin) for overnight incubation at 37°C under an anaerobic condition. After 20-h incubation, fluid medium is removed and added with 2 ml of pre-warmed, fresh THYE plus 5% horse serum. The cultures are incubated for 30 minutes and each well is supplemented with a final concentration of 200 ng/ml of synthetic competence stimulating peptide (SCSP) and varying concentrations of the inhibitor and the incubation is continued. After 30 minutes, plasmid DNA (1 mg/ml) or chromosomal DNA (10 mg/ml) is added to each well and the cultures are incubated for an additional 2 hr. Planktonic cells are then removed and the wells are washed once with PBS buffer. Biofilm cells are collected into 2 ml fresh medium by a gentle sonication or washing the wells using a pipette. The samples are centrifuged at 12,000 x g for 5 min. Both biofilm and planktonic cells are resuspended into 200 μl of fresh medium and are immediately spread on THYE agar plus appropriate antibiotics. Transformation frequency is determined after 48-h of incubation.

[0086] Homologues of the *Streptococcus pneumoniae comD* [SEQ ID NO:5]/*E* [SEQ ID NO:7] genes encoding a histidine kinase [SEQ ID NO:6]/ response regulator [SEQ ID NO:8] system were identified. This sequence was used to design primers to amplify the region from a number of *S. mutans* isolates. An open reading frame consisting of 138 nucleotides was located 148 nucleotides 5' proximal from the end of the *comD* homolog in the opposite orientation (Fig 1). This ORF was found to encode a peptide of 46-amino acid [SEQ ID NO:2] in length, the precursor of the 21-amino acid CSP [SEQ ID NO:4].

[0087] The *comCDE* genes [SEQ ID NO:23] were amplified from the genomes of several *S. mutans* isolates by PCR using primers designed based on the genome database sequence and their nucleotide sequences determined. The deduced amino acid sequences are compared among the isolates by sequence alignment to confirm identity.

[0088] Genes are inactivated by integration of internal homologous fragments into the suicide vector pVA8912. Mutants defective in each of the individual genes (*comC* [SEQ ID NO:1], *comD* [SEQ ID NO:5], *comE* [SEQ ID NO:7]) are inactivated and their phenotypes are compared to the parent strain NG8 for their abilities to form biofilms, tolerate acidic pH (pH 2-4), and transport and incorporate DNA. The knockout mutants of *com D* [SEQ ID NO:5] and *E* [SEQ ID NO:7] were constructed by insertion-duplication mutagenesis, whereas the knockout *comC* [SEQ ID NO:1] mutant was created by allelic exchange via insertion of an erythromycin resistance determinant into the *comC* [SEQ ID NO:1] locus (Li *et al*, 2001). All mutant strains were therefore resistant to erythromycin. The wild-type strain was subcultured routinely on

Todd-Hewitt-Yeast Extract (THYE) agar plates (BBL®; Becton Dickinson, Cockeysville, MD), whereas the mutants were maintained on THYE agar plus 10 μg/ml of erythromycin. A minimal medium (DMM) was prepared to grow biofilms by a modification of the method described previously (Loo *et al*, 2000). The medium contained 58 mM K_2HPO_4, 15 mM KH_2PO_4, 10 mM $(NH_4)_2SO_4$, 35 mM NaCl, 2 mM $MgSO_2 \cdot 7H_2O$, 0.2% (wt/vol) Casamino Acids and was supplemented with filter-sterilized vitamins, (0.04 mM nicotinic acid, 0.1 mM pyridoxine HCl, 0.01 mM pantothenic acid, 1 μM riboflavin, 0.3 μM thiamin HCl, and 0.05 μM D-biotin), amino acids (4 mM L-glutamic acid, 1 mM L-arginine HCl, 1.3 mM L-cysteine HCl, and 0.1 mM L-tryptophan) and 20 mM glucose.

[0090] To determine if the synthetic peptide **[SEQ ID NO:16]** could restore defective phenotypes of the *comC* **[SEQ ID NO:2]** mutants, a chemically synthesized 21-amino acid competence-stimulating peptide (CSP) **[SEQ ID NO:4]** (Li et al, 2001) was used in complementary experiments. The peptide was freshly dissolved in sterile distilled water to a concentration of 1 mg/ml. The CSP solution was then added to the cultures at a final concentration of 2 μg/ml 2 h after inoculation of bacterial cells.

[0096] Two methods modified based on the protocols described by Perry *et al* (Infect Immun, 41:722-727) and Lindler and Macrina (J Bacteriol, 166:658-665) were used to assay natural transformation of biofilm cells. Biofilms formed on polystyrene microtiter plates were added with 2 ml of pre-warmed, fresh THYE plus 5% horse serum (THYE-HS) immediately following removal of the BM medium, and the incubation continued at 37^0C. After 2h incubation, a final concentration of 1 μg/ml plasmid DNA or 10 μg/ml of chromosomal DNA was added to each well. The cultures were incubated for an additional 2 h before collection of the cells for plating. To assay induction of competence by synthetic competence stimulating peptide (SCSP) **[SEQ ID NO:16]**, the cultures were incubated for 30 min and a final concentration of 500 ng/ml of SCSP) **[SEQ ID NO:16]** was added to each well. After a 30 min incubation, equal amounts of DNA was added to each well (1 μg/ml plasmid or 10 μg/ml of chromosomal DNA) and incubation continued for another 2 h. Fluid medium was then removed from individual wells and the wells were washed once with PBS buffer. Biofilm cells were collected into 2 ml PBS buffer by gentle sonication or by washing the wells using a pipette. The samples were centrifuged at 12,000 X g for 5 min. Both biofilm and planktonic cells were resuspended into 200 μl of fresh medium and were immediately spread on THYE agar plates plus appropriate antibiotics. For the biofilms developed in the chemostat, rods with biofilm cells were removed and placed into 2 ml of pre-warmed, fresh THYE-HS medium for 30 min incubation. Transformation was then

initiated by using the same methods as described above. The planktonic cells were also removed to compare the transformation frequency. After completion of the transformation procedures, both biofilm and planktonic cells were spread on THYE agar plus appropriate antibiotic. Transformation frequency was assessed after 48-h incubation. Transformation frequency was expressed as the number of transformants per μg DNA per viable recipient at the time of DNA added.

In the Claims:

3. (Amended) The compound of claim 2, comprising a derivative of [SEQ ID NO:2] **or [SEQ ID NO:4]**, a fragment of [SEQ ID NO:2] **or [SEQ ID NO:4]** or a derivative of a fragment of [SEQ ID NO:2] **or [SEQ ID NO:4]**.

4. (Amended) The compound of claim 3, wherein amino acids are removed from the N-terminus and/or C-terminus of [SEQ ID NO:2] **or [SEQ ID NO:4]**.

5. (Amended) A pharmaceutical composition comprising all or part of the peptide of ~~claims 1 to 4~~ **claim 1** and a carrier.

6. (Amended) A method of medical treatment or prophylaxis of caries or endocarditis, comprising administering the compound of **claim 1**~~any of claims 1 to 4 or the pharmaceutical composition of claim 5~~.

7. CANCELLED

8. CANCELLED

9. CANCELLED

10. CANCELLED

11. CANCELLED

12. CANCELLED

13. CANCELLED

14. CANCELLED

15. CANCELLED

16. CANCELLED

17. CANCELLED

18. CANCELLED

19. CANCELLED

20. CANCELLED

21. CANCELLED

24. (Amended) The polypeptide of claim 23 comprising all or part of an amino acid sequence in [SEQ ID NO:2] **or [SEQ ID NO:4]**.

29. (Amended) An isolated nucleic acid molecule encoding the polypeptide of any of claims ~~21~~ **22** to 28.

30. (Amended) An antibody directed against the polypeptide of any of claims ~~21~~ **22** to 28.

32. (Amended) A vaccine composition comprising all or part of the peptide of any of claims ~~21~~ **22** to 28 and a carrier.

34. CANCELLED

35. CANCELLED

36. CANCELLED

37. CANCELLED

APPENDIX - 2

COMPLETE SET OF PENDING CLAIMS, AS AMENDED ON SEPTEMBER 17, 2001

1. A compound that competitively inhibits binding of CSP to *S. mutans* histidine kinase.

2. The compound of claim 1, comprising a peptide or an antibody.

3. (Amended) The compound of claim 2, comprising a derivative of [SEQ ID NO:2] or [SEQ ID NO:4], a fragment of [SEQ ID NO:2] or [SEQ ID NO:4] or a derivative of a fragment of [SEQ ID NO:2] or [SEQ ID NO:4].

4. (Amended) The compound of claim 3, wherein amino acids are removed from the N-terminus and/or C-terminus of [SEQ ID NO:2] or [SEQ ID NO:4].

5. (Amended) A pharmaceutical composition comprising all or part of the peptide of claim 1 and a carrier.

6. (Amended) A method of medical treatment or prophylaxis of caries or endocarditis, comprising administering the compound of claim 1.

22. An isolated CSP or a fragment thereof having *S. mutans* CSP activity.

23. The polypeptide of claim 22 comprising a *S. mutans* CSP.

24. (Amended) The polypeptide of claim 23 comprising all or part of an amino acid sequence in [SEQ ID NO:2] or [SEQ ID NO:4].

25. A polypeptide fragment of the peptide of claim 24, or a peptide mimetic of the CSP.

26. The polypeptide of claim 24 which is recombinantly produced.

27. A polypeptide comprising a sequence having greater than 30%, 50% or 60% sequence identity to the polypeptide of claim 24.

28. The polypeptide of claim 24, isolated from *S. mutans*.

29. (Amended) An isolated nucleic acid molecule encoding the polypeptide of any of claims 22 to 28.

30. (Amended) An antibody directed against the polypeptide of any of claims 22 to 28.

31. The antibody of claim 30, comprising a monoclonal antibody or a polyclonal antibody.

32. (Amended) A vaccine composition comprising all or part of the peptide of any of claims 22 to 28 and a carrier.

33. The vaccine composition of claim 32, wherein the peptide is coupled to a compound comprising all or part of KLH, ovalbumin, or thyroglobulin.

APPENDIX - 3

DETAILS OF DRAWING REVISIONS SUBMITTED ON SEPTEMBER 17, 2001

Figure 1

~~Schematic layout of the arrangement of the genetic locus encoding the signal peptide precursor, the histidine kinase and the response regulator. Note that this arrangement is different from other loci in related streptococci for the following reasons: a) The comC gene is transcribed from its own promoter alone, unlike the genes thus far described in other streptococci that are arranged in an operon-like cluster with the comC/DE genes being transcribed from a single promoter.~~
~~b) The comC gene is separated from the *comD* gene by 148 nucleotides.~~

Streptococcus mutans
***ComCDE* Operon**

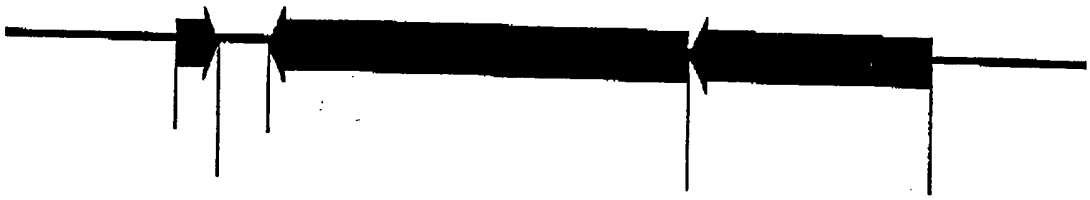

Figure 2

~~Sequences of the open reading frames encoding the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).~~

(A.) ~~S. mutans comC gene~~

Encodes a precursor to a signal peptide

[ATGAAAAAAACACTATCATTAAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAG
AGATTATCATTGGCGGA(AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAG
TTTTACACAAGCTTTGGGAAAA)]TAA

(B.) ~~S. mutans CSP encoding sequence~~

~~Competence Signal Peptide~~

AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAGTTTTACACAAGCTTTGGG
AAAA [SEQ ID NO:1]

(C.) ~~S. mutans comD gene~~

~~Encodes a protein that functions as a histidine kinase receptor~~

[ATGAATGAAGCCTTAATGATACTTTCAAATGGTTTATTAACTTATCTAACCGTTCTAT
TTCTCTTGTTTCTATTTTCTAAGGTAAGTAATGTCACTTTATCGAAAAAGGAATTAACT
CTTTTTTCGATAAGCAATTTTCTGATAATGATTGCTGTTACGATGGTGAACGTAAACCT
GTTTTATCCTGCAGAGCCTCTTTATTTTATAGCTTTATCAATTTATCTTAATAGACAGA
ATAGTCTTTCTCTAAATATATTTTATGGTCTGCTGCCTGTTGCCAGTTCTGACTTGTTT
AGGCGGGCAATCATATTCTTTATCTTGGATGGAACTCAAGGAATTGTAATGGGCAGTAG
CATTATAACCACCTATATGATCGAGTTTGCAGGAATAGCGCTAAGTTACCTCTTTCTCA
GTGTGTTCAATGTTGATATTGGTCGACTTAAAGATAGTTTGACCAAGATGAAGGTCAAA
AAACGCTTGATTCCAATGAATATTACTATGCTTCTATACTACCTTTTAATACAGGTATT
GTATGTTATAGAGAGTTATAATGTGATACCGACTTTAAAATTTCGTAAATTTGTCGTTA
TTGTCTATCTTATTTTATTTTTGATTCTGATCTCATTTTTAAGCCAATATACCAAACAA
AAGGTTCAAAATGAGATAATGGCACAAAAGGAAGCTCAGATTCGAAATATCACCCAGTA
TAGTCAGCAAATAGAATCTCTTTACAAGGATATTCGAAGTTTCCGCCATGATTATCTGA
ATATTTTAACTAGCCTCAGATTAGGCATTGAAAATAAAGATTTAGCTAGTATTGAAAAG
ATTTACCATCAAATCTTAGAAAAAACAGGACATCAATTGCAGGATACCCGTTATAATAT
CGGCCATCTAGCTAATATTCAAAACGATGCTGTCAAGGGTATCTTGTCAGCAAAAATCT
TAGAAGCTCAGAATAAAAAGATTGCTGTCAATGTAGAAGTCTCAAGTAAAATACAACTG
CCTGAGATGGAGTTGCTTGATTTCATTACCATACTTTCTATCTTGTGTGATAATGCCAT
TGAGGCTGCTTTCGAATCATTAAATCCTGAAATTCAGTTAGCCTTTTTTAAGAAAAATG
GCAGTATAGTCTTTATCATTCAGAATTCCACCAAAGAAAAACAAATAGATGTGAGTAAA
ATTTTTAAAGAAAACTATTCCACTAAAGGCTCCAATCGCGGTATTGGTTTAGCAAAGGT
GAATCATATTCTTGAACATTATCCCAAAACCAGTTTACAAACAAGCAATCATCATCATT
TATTCAAGCAACTCCTAATAATAAAA] TAG

(D.) ~~S. mutans comE gene~~

~~Encodes a response regulator that activates transcription of a number of genes~~

Figure 2 (cont'd)

[ATGATTTCTATTTTTGTATTGGAAGATGATTTTTTACAACAAGGACGTCTTGAAACCA
CCATTGCAGCTATCATGAAAGAAAAAAATTGGTCTTATAAAGAATTGACTATTTTTGGA
AAACCACAACAACTTATTGACGCTATCCCTGAAAAGGGCAATCACCAGATTTTCTTTTT
GGATATTGAAATCAAAAAAGAGGAAAAGAAAGGACTGGAAGTAGCCAATCAGATTAGAC
AGCATAATCCTAGTGCAGTTATTGTCTTTGTCACGACACATTCTGAGTTTATGCCCCTC
ACTTTTCAGTATCAGGTATCTGCTTTGGATTTTATTGATAAATCTTTGAATCCTGAGGA
GTTCTCCCACCGCATTGAATCAGCGCTGTATTATGCTATGGAAAACAGCCAGAAGAATG
GTCAATCAGAGGAACTTTTTATTTTCCATTCATCTGAAACTCAGTTTCAGGTCCCTTTT
GCTGAGATTCTGTATTTTGAAACATCTTCAACAGCCCATAAGCTCTGCCTTTATACTTA
TGATGAACGGATTGAATTCTACGGCAGTATGACTGACATTGTTAAAATGGATAAGAGAC
TTTTTCAGTGCCATCGCTCTTTTATTGTCAATCCTGCCAATATTACCCGTATTGATCGG
AAAAAACGCTTGGCCTATTTTCGAAATAATAAGTCTTGTCTTATTTCACGAACTAAGTT
AACAAAACTGAGAGCTGTGATTGCTGATCAAAGGAGAGCAAAA] TGA

Figure 3

~~The amino acid sequences of the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).~~

A.

~~S. mutans ComC protein (CSP Precursor)~~

MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK

B.

~~S. mutans ComD protein (Histidine Kinase)~~

MNEALMILSNGLLTYLTVLFLLFLFSKVSNVTLSKKELTLFSISNFLIMIAVTMVNVNL
FYPAEPLYFIALSIYLNRQNSLSLNIFYGLLPVASSDLFRRAIIFFILDGTQGIVMGSS
IITTYMIEFAGIALSYLFLSVFNVDIGRLKDSLTKMKVKKRLIPMNITMLLYYLLIQVL
YVIESYNVIPTLKFRKFVVIVYLILFLILISFLSQYTKQKVQNEIMAQKEAQIRNITQY
SQQIESLYKDIRSFRHDYLNILTSLRLGIENKDLASIEKIYHQILEKTGHQLQDTRYNI
GHLANIQNDAVKGILSAKILEAQNKKIAVNVEVSSKIQLPEMELLDFITILSILCDNAI
EAAFESLNPEIQLAFFKKNGSIVFIIQNSTKEKQIDVSKIFKENYSTKGSNRGIGLAKV
NHILEHYPKTSLQTSNHHHLFKQLLIIK

C.

~~S. mutans ComE protein (Response Regulator)~~

MISIFVLEDDFLQQGRLETTIAAIMKEKNWSYKELTIFGKPQQLIDAIPEKGNHQIFFL
DIEIKKEEKKGLEVANQIRQHNPSAVIVFVTTHSEFMPLTFQYQVSALDFIDKSLNPEE
FSHRIESALYYAMENSQKNGQSEELFIFHSSETQFQVPFAEILYFETSSTAHKLCLYTY
DERIEFYGSMTDIVKMDKRLFQCHRSFIVNPANITRIDRKKRLAYFRNNKSCLISRTKL
TKLRAVIADQRRAK

Figure 4

~~The deduced amino acid sequence of the signal peptide precursor in various strains and its predicted cleavage site. The original peptide is expressed as a 46 amino acid peptide that is cleaved after the glycine-glycine residues to generate an active signal peptide.~~

```
BM71 CSP      1 MKKTPSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
GB14 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
H7 CSP        1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
JH1005 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGTLSTFFRLFNRSFTQA    43
LT11 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
NG8 CSP       1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
UAB159 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
                **** ***********************.***************

consensus: 1 MKKTLSLKNDFKEIKTDELEIIIGG SGSLSTFFRLFNRSFTQALGK 46
predicted cleavage site:              ^
```

Figure 5

~~The synthetic signal peptide that is effective at inducing competence, biofilm formation and acid tolerance in *Streptococcus mutans*.~~

SGSLSTFFRLFNRSFTQALGK [SEQ ID NO:2]

Figure 6
~~The natural activity of the signal/receptor system functioning *in vitro* in model biofilms as determined by the ability of various strains of *S. mutans* to accept donor plasmid DNA conferring erythromycin resistance.~~

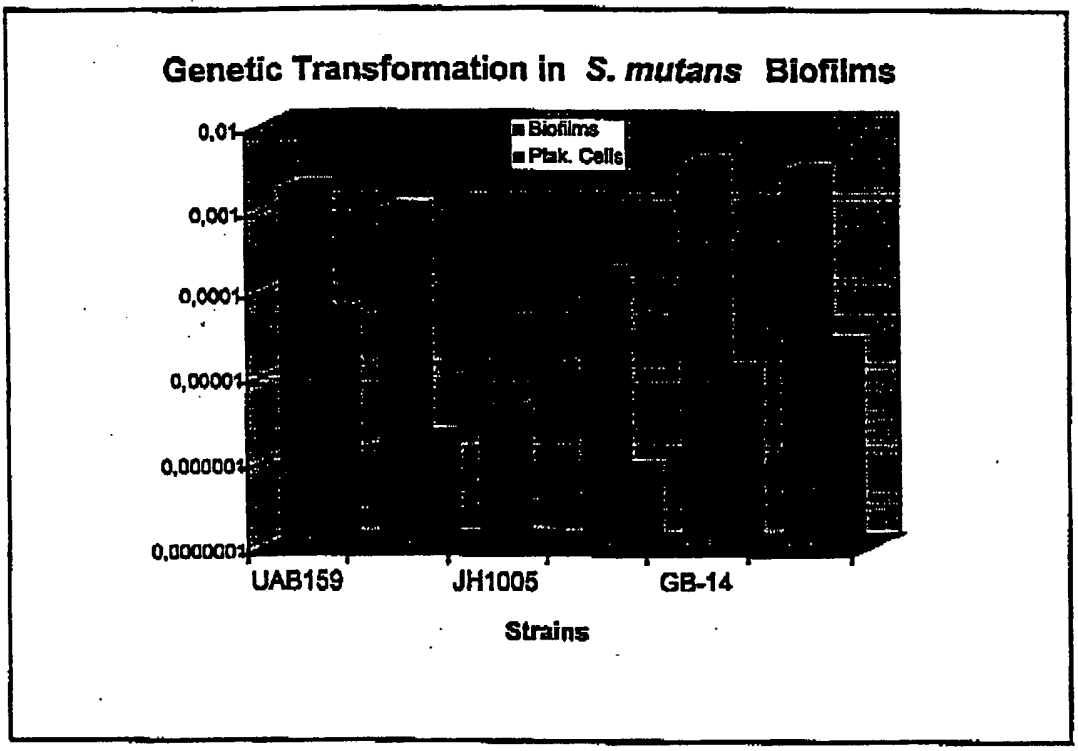

Figure 7
~~Induction of genetic transformation in *Streptococcus mutans* by synthetic competence stimulating peptide (SCSP)[1]~~

Strain	Peptide added Number of Transformants/Recipients	No peptide Number of Transformants/Recipients
UAB15	4.65×10^{-1}	1.78×10^{-6}
JH1005[2]	6.98×10^{-2}	0

[1]The final concentration of SCSP used was 500 ng/ml.
The strain contains a nonsense mutation in the *comC* gene encoding the CSP.

Figure 8

~~List of the primers used to amplify the genes or internal regions of the target genes by polymerase chain reaction (PCR) for subsequent sequencing or inactivation.~~

ComC region

ComC Primer Pair: F5-B5

[F5] 23406-23424 5'- AGTTTTTTGTCTGGCTGCG -3'
19 nt forward primer
pct G+C: 47.4 Tm: 50.5

[B5] 24056-24037 5'- TCCACTAAAGGCTCCAATCG -3'
20 nt backward primer
pct G+C: 50.0 Tm: 51.9

651 nt product for F5-B5 pair (23406-24056)
Optimal annealing temp: 50.3
pct G+C: 30.9 Tm: 71.5

ComD region

ComD Primer Pair: F1-B1

[F1] 392-415 5'- CGCTAAGTTACCTCTTTCTCAGTG -3'
24 nt forward primer
pct G+C: 45.8 Tm: 51.6

[B1] 683-663 5'- GCTTCCTTTTGTGCCATTATC -3'
21 nt backward primer
pct G+C: 42.9 Tm: 50.8

292 nt product for F1-B1 pair (392-683)
Optimal annealing temp: 49.5
pct G+C: 30.8 Tm: 70.2

ComE region

ComE Primer Pair: F1-B1

[F1] 145-165 5'- CCTGAAAAGGGCAATCACCAG -3'
21 nt forward primer
pct G+C: 52.4 Tm: 55.9

[B1] 606-585 5'- GCGATGGCACTGAAAAAGTCTC -3'
22 nt backward primer
pct G+C: 50.0 Tm: 55.4

462 nt product for F1-B1 pair (145-606)
Optimal annealing temp: 53.6
pct G+C: 38.3 Tm: 74.1

Figure 9

~~ComCDE local region. The ComC (first highlighted region; nucleotides 101 to 241), ComD (second highlighted region; nucleotides 383 to 1708) and ComE (third highlighted region; nucleotides 1705 to 2457) proteins are highlighted.~~

Sequence Range: 1 to 2557

```
          10         20         30         40         50
ACATTATGTGTCCTAAGGAAAATATTACTTTTTCAAGAAAATCCATGATT
TGTAATACACAGGATTCCTTTTATAATGAAAAAGTTCTTTTAGGTACTAA
                                  <K  K  L  F  I  W  S  K
                                 <_____________________

          60         70         80         90         100
TTTTCATAAAAAAATAGTATACTAATTATAATCAAAAAAAGGAGATATAAA
AAAAGTATTTTTTATCATATGATTAATATTAGTTTTTTTCCTCTATATTT
  <K  M  F  F  L  I  S  I  I  I  L  F  L  L  Y  L
 <_____________________________________________________

          110        120        130        140        150
ATGAAAAAAACACTATCATTAAAAAATGACTTTAAAGAAATTAAGACTGA
TACTTTTTTTGTGATAGTAATTTTTTACTGAAATTTCTTTAATTCTGACT
  M  K  K  T  L  S  L  K  N  D  F  K  E  I  K  T  D>
_____________________ORF RF[2] _____________________>
<I  F  F  V  S  D  N  F  F  S  K  L  S  I  L  V  S
 <_____________________________________________________

          160        170        180        190        200
TGAATTAGAGATTATCATTGGCGGAAGCGGAAGCCTATCAACATTTTTCC
ACTTAATCTCTAATAGTAACCGCCTTCGCCTTCGGATAGTTGTAAAAAGG
  E  L  E  I  Y  I  G  G  S  G  S  L  S  T  F  P>
_____________________ORF RF[2] _____________________>
<S  N  S  I  I  M
 <_____________

          210        220        230        240        250
GGCTGTTTAACAGAAGTTTTACACAAGCTTTGGGAAAATAAGATAGGCTA
CCGACAAATTGTCTTCAAAATGTGTTCGAAACCCTTTTATTCTATCCGAT
K  L  F  N  R  S  F  T  Q  A  L  G  K>
_____________ORF RF[2] _____________>

          260        270        280        290        300
ACATTGGAATAAAACAAGGCTGGATTTATTATTCCAGCCTTTTTAAATGT
TGTAACCTTATTTTGTTCCGACCTAAATAATAAGGTCGGAAAAATTTACA

          310        320        330        340        350
AAAATAAAAATACAGGGTTAAATAATCAAGTGTGCTGTCGTGGATGAGAA
TTTTATTTTTATGTCCCAATTTATTAGTTCACACGACAGCACCTACTCTT

          360        370        380        390        400
GATAAAACTATCTCTTAGAGAATAGGCCTCCTCTATTTTATTATTAGGAG
CTATTTTGATAGAGAATCTCTTATCCGGAGGAGATAAAATAATAATCCTC
                                  <K  I  L  L  L
                                 <____ORF RF[____

          410        420        430        440        450
```

Figure 9 (cont'd-1)

```
AACGAACTTATTTACTACTACTAACGAACAAACATTTGACCAAAACCCTA
 Q  K  P  L  H  H  R  N  S  T  Q  I  S  T  K  P  T
<_____________________ORF RF[4] C_____________________

       460        470        480        490        500
AATGTTCAAGAATATGATTCACCTTTGCTAAACCAATACCGCGATTGGAG
TTACAAGTTCTTATACTAAGTGGAAACGATTTGGTTATGGCGCTAACCTC
 <H  E  L  I  H  N  V  K  A  L  G  I  G  R  N  S
<_____________________ORF RF[4] C_____________________

       510        520        530        540        550
CCTTTAGTGGAATAGTTTTCTTTAAAAATTTTACTCACATCTATTTGTTT
GGAAATCACCTTATCAAAAGAAATTTTTAAAATGAGTGTAGATAAACAAA
<G  K  T  S  V  N  E  K  F  I  K  S  V  D  I  Q  K
<_____________________ORF RF[4] C_____________________

       560        570        580        590        600
TTCTTTGGTGGAATTCTGAATGATAAAGACTATACTGCCATTTTTCTTAA
AAGAAACCACCTTAAGACTTACTATTTCTGATATGACGGTAAAAAGAATT
<E  K  T  S  N  Q  I  I  F  V  I  S  G  N  K  K  F
<_____________________ORF RF[4] C_____________________

       610        620        630        640        650
AAAAGGCTAACTGAATTTCAGGATTTAATGATTCGAAAGCAGCCTCAATG
TTTTCCGATTGACTTAAAGTCCTAAATTACTAAGCTTTCGTCGGAGTTAC
                                                   M>
                                                    >
 <F  A  L  Q  I  E  P  N  L  S  E  F  A  A  E  I
<_____________________ORF RF[4] C_____________________

       660        670        680        690        700
GCATTATCACACAAGATAGAAAGTATGGTAATGAAATCAAGCAACTCCAT
CGTAATAGTGTGTTCTATCTTTCATACCATTACTTTAGTTCGTTGAGGTA
 A  L  S  H  K  I  E  S  M  V  M  K  S  S  N  S  I>
                    ORF RF[3]                      >
<A  N  D  C  E  I  S  L  I  T  I  F  D  L  E  E  M
<_____________________ORF RF[4] C_____________________

       710        720        730        740        750
CTCAGGCAGTTGTATTTTACTTGAGACTTCTACATTGACAGCAATCTTTT
GAGTCCGTCAACATAAAATGAACTCTGAAGATGTAACTGTCGTTAGAAAA
  S  G  S  C  I  L  L  E  T  S  T  L  T  A  I  F>
                    ORF RF[3]                      >
<E  P  L  Q  I  K  S  S  V  E  V  N  V  A  I  K  K
<_____________________ORF RF[4] C_____________________

       760        770        780        790        800
TATTCTGAGCTTCTAAGATTTTTGCTGACAAGATACCCTTGACAGCATCG
ATAAGACTCGAAGATTCTAAAAACGACTGTTCTATGGGAACTGTCGTAGC
L  F>
    >
 <N  Q  A  E  L  I  K  A  S  L  I  G  K  V  A  D
<_____________________ORF RF[4] C_____________________

       810        820        830        840        850
TTTTGAATATTAGCTAGATGGCCGATATTATAACGGGTATCCTGCAATTG
AAAACTTATAATCGATCTACCGGCTATAATATTGCCCATAGGACGTTAAC
```

Figure 9 (cont'd-2)

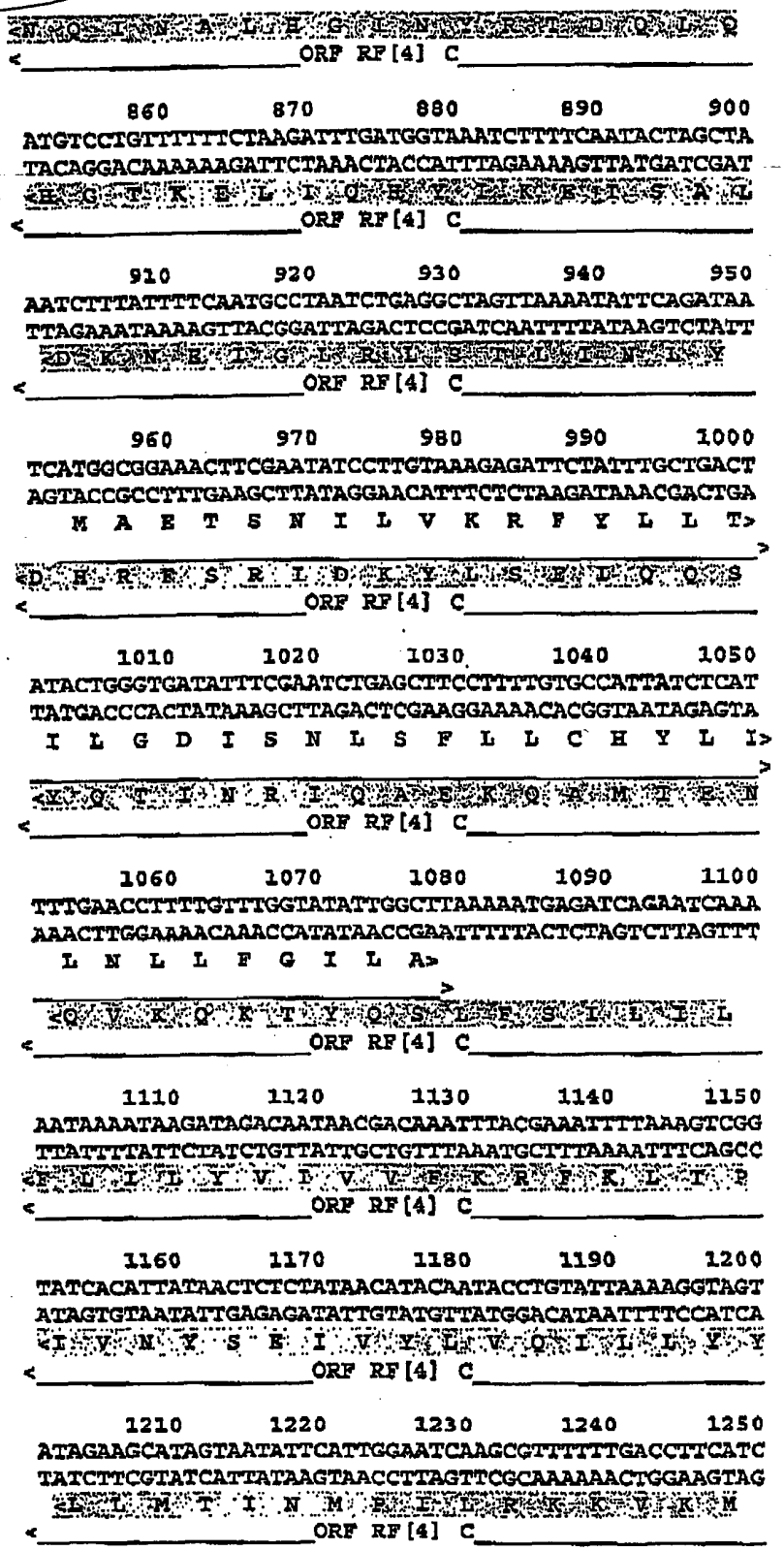

Figure 9 (contd-3)

```
      1260       1270       1280       1290       1300
TTGGTCAAACTATCTTTAAGTCGACCAATATCAACATTGAACACACTGAG
AACCAGTTTGATAGAAATTCAGCTGGTTATAGTTGTAACTTGTGTGACTC
<K  P  L  S  D  K  L  R  Q  L  D  V  N  F  V  S  L
<_____________________ORF RF[4] C_____________________

      1310       1320       1330       1340       1350
AAAGAGGTAACTTAGCGCTATTCCTGCAAACTCGATCATATAGGTGGTTA
TTTCTCCATTGAATCGCGATAAGGACGTTTGAGCTAGTATATCCACCAAT
<F  L  K  S  L  A  I  G  A  K  P  I  M  Y  D  T  I
<_____________________ORF RF[4] C_____________________

      1360       1370       1380       1390       1400
TAATGCTACTGCCCATTACAATTCCTTGAGTTCCATCCAAGATAAAGAAT
ATTACGATGACGGGTAATGTTAAGGAACTCAAGGTAGGTTCTATTTCTTA
<I  S  S  G  M  V  I  G  Q  T  G  D  L  I  F  F
<_____________________ORF RF[4] C_____________________
                      <L  E  K  L  E  M  W  S  L  S  Y
                      <_____________________

      1410       1420       1430       1440       1450
ATGATTGCCCGCCTAAACAAGTCAGAACTGGCAACAGGCAGCAGACCATA
TACTAACGGGCGGATTTGTTCAGTCTTGACCGTTGTCCGTCGTCTGGTAT
<I  I  A  R  R  N  K  L  D  S  W  S  A  V  P  I  G  Y
<_____________________ORF RF[4] C_____________________
  <S  Q  G  G  L  C  T  L  V  P  L  L  C  C  V  M
<_____________________________________________

      1460       1470       1480       1490       1500
AAATATATTTAGAGAAAGACTATTCTGTCTATTAAGATAAATTGATAAAG
TTTATATAAATCTCTTTCTGATAAGACAGATAATTCTATTTAACTATTTC
<F  I  N  L  S  L  S  N  Q  R  N  L  Y  L  S  L  A
<_____________________ORF RF[4] C_____________________

      1510       1520       1530       1540       1550
CTATAAAATAAAGAGGCTCTGCAGGATAAAACAGGTTTACGTTCACCATC
GATATTTTATTTCTCCGAGACGTCCTATTTTGTCCAAATGCAAGTGGTAG
  <I  K  Y  L  P  E  A  P  V  F  L  N  V  N  V  M
<_____________________ORF RF[4] C_____________________

      1560       1570       1580       1590       1600
GTAACAGCAATCATTATCAGAAAATTGCTTATCGAAAAAAGAGTTAATTC
CATTGTCGTTAGTAATAGTCTTTTAACGAATAGCTTTTTTCTCAATTAAG
<T  V  A  I  M  I  L  F  N  S  I  S  F  L  I  E  E
<_____________________ORF RF[4] C_____________________

      1610       1620       1630       1640       1650
CTTTTTCGATAAAGTGACATTACTTACCTTAGAAAATAGAAACAAGAGAA
GAAAAAGCTATTTCACTGTAATGAATGGAATCTTTTATCTTTGTTCTCTT
<K  K  S  L  T  V  N  S  V  K  S  F  L  F  L  L  F
<_____________________ORF RF[4] C_____________________

      1660       1670       1680       1690       1700
ATAGAACGGTTAGATAAGTTAATAAACCATTTGAAAGTATCATTAAGGCT
TATCTTGCCAATCTATTCAATTATTTGGTAAACTTTCATAGTAATTCCGA
  <L  V  T  L  Y  T  L  L  G  N  S  L  I  M  L  A
<_____________________ORF RF[4] C_____________________
```

Figure 9 (cont'd-4)

```
      1710       1720       1730       1740       1750
TCATTCATTTTGCTCTCCTTTGATCAGCAATCACAGCTCTCAGTTTTGTT
AGTAAGTAAAACGAGAGGAAACTAGTCGTTAGTGTCGAGAGTCAAAACAA
<E  N  M
<_________
      <K  A  R  R  Q  D  A  I  V  A  R  L  K  T
     <_____________ORF RF[5] C_____________

      1760       1770       1780       1790       1800
AACTTAGTTCGTGAAATAAGACAAGACTTATTATTTCGAAAATAGGCCAA
TTGAATCAAGCACTTTATTCTGTTCTGAATAATAAAGCTTTTATCCGGTT
<L  K  T  R  S  L  L  C  S  K  N  N  R  F  Y  A  L
<_____________ORF RF[5] C_____________

      1810       1820       1830       1840       1850
GCGTTTTTTCCGATCAATACGGGTAATATTGGCAGGATTGACAATAAAAG
CGCAAAAAAGGCTAGTTATGCCCATTATAACCGTCCTAACTGTTATTTTC
<R  K  K  R  D  L  R  T  I  N  A  P  D  V  I  F  S
<_____________ORF RF[5] C_____________

      1860       1870       1880       1890       1900
AGCGATGGCACTGAAAAAGTCTCTTATCCATTTTAACAATGTCAGTCATA
TCGCTACCGTGACTTTTTCAGAGAATAGGTAAAATTGTTACAGTCAGTAT
     M  A  L  K  K  S  L  I  H  F  N  N  V  S  H>
     _____________ORF RF[1] _____________>
  <R  H  C  Q  F  L  R  K  D  N  K  V  I  D  T  M
<_____________ORF RF[5] C_____________
                                          <V
                                          <__

      1910       1920       1930       1940       1950
CTGCCGTAGAATTCAATCCGTTCATCATAAGTATAAAGGCAGAGCTTATG
GACGGCATCTTAAGTTAGGCAAGTAGTATTCATATTTCCGTCTCGAATAC
T  A  V  E  F  N  P  F  I  I  S  I  K  A  E  L  M>
_____________ORF RF[1] _____________>
<S  G  Y  F  E  L  R  E  D  V  T  Y  Y  C  I  R  H
<_____________ORF RF[5] C_____________
  <A  T  S  N  L  G  N  M  M  L  I  F  A  S  S  I
<_____________ORF RF[6] C_____________

      1960       1970       1980       1990       2000
GGCTGTTGAAGATGTTTCAAAATACAGAATCTCAGCAAAAGGGACCTGAA
CCGACAACTTCTACAAAGTTTTATGTCTTAGAGTCGTTTTCCCTGGACTT
 G  C>
_____>
 <A  T  S  S  T  E  F  Y  L  I  E  A  F  P  V  Q  F
<_____________ORF RF[5] C_____________
<P  Q  Q  L  H  K  L  I  C  F  R  L  L  L  S  R  F
<_____________ORF RF[6] C_____________

      2010       2020       2030       2040       2050
ACTGAGTTTCAGATGAATGGAAAATAAAAAGTTCCTCTGATTGACCATTC
TGACTCAAAGTCTACTTACCTTTTATTTTTCAAGGAGACTAACTGGTAAG
  Q  T  E  S  S  H  F  I  F  L  E  E  S  Q  G  N
<_____________ORF RF[5] C_____________
<S  L  K  L  H  I  S  F  L  F  N  R  Q  N  V  M  R
```

Figure 9 (cont'd-5)

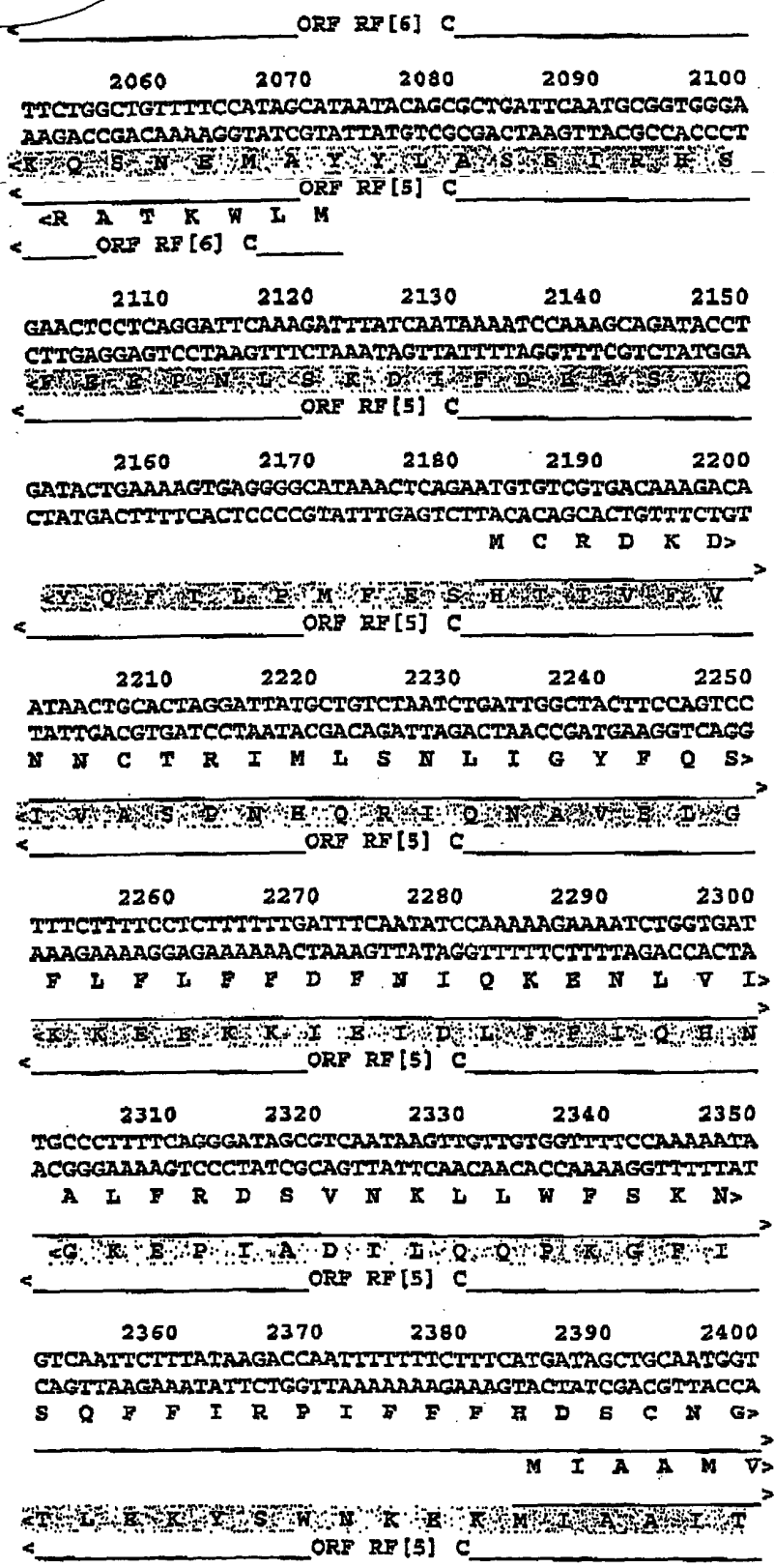

Figure 9 (cont'd-6)

```
      2410       2420       2430       2440       2450
GGTTTCAAGACGTCCTTGTTGTAAAAAATCATCTTCCAATACAAAAATAG
CCAAAGTTCTGCAGGAACAACATTTTTTAGTAGAAGGTTATGTTTTTATC
 G  F  K  T  S  L  L>
 _____________>
  V  S  R  R  P  C  C  K  K  S  S  S  N  T  K  I>
 _____________________________________________>
<T  E  L  R  G  Q  Q  L  R  D  D  R  L  V  R  I  S
<_____________________ORF RF[5] C_____________________

      2460       2470       2480       2490       2500
AAATCATTATTTCTCCTTTAATCTTCTATTTAGGTTAGCTGATTAACACT
TTTAGTAATAAAGAGGAAATTAGAAGATAAATCCAATCGACTAATTGTGA
E  I  I  I  S  P  L  I  F  Y  L  G>
 _____________________________________>
  <I  M
<_____

      2510       2520       2530       2540       2550
ATACACAGAAAAGGTATAAAACGATATCACTCAATAAAATCTACTAACTT
TATGTGTCTTTTCCATATTTTGCTATAGTGAGTTATTTTAGATGATTGAA

AATAACC
TTATTGG
```

Figure 10

~~The comX nucleotid sequence, amino acid sequence, and its local region with 100bp included both upstream and downstream (promoter is upstream).~~

A.

~~> S. mutans comX gene~~

ATGGAAGAAGATTTTGAAATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAG
CCGTTATTACTTTATTAAAATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGA
TTTTGCACCAATTATTAAGGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTAT
ATCTATTTTAAGACACGTTTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAG
TCAGAAACGTCGTTTTAATAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTT
TGTCAAGTGGCGGTATGCAATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCA
TATAAACAAGGTCTGAGTACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGA
GCACTTTTTGGGAAGGCAAAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTA
AGGAAAAA

B.

~~> S. mutans ComX protein~~

MEEDFEIVFNKVKPIVWKLSRYYFIKMWTREDWQQEGMLILHQLLREHPELEEDDTKLY
IYFKTRFSNYIKDVLRQQESQKRRFNRMSYEEVGEIEHCLSSGGMQLDEYILFRDSLLA
YKQGLSTEKQELFERLVAGEHFLGRQSMLKDLRKKLSDFKEK

C.

~~> S. mutans comX gene local region~~

GTAAATAAAACAGCCAGTTAAGATGGGACATTTATGTCCTGTTCTTAAAGTCTTTTTCG
TTTTATAATAATTTTATTATAAAAGGAGGTCATCGTAATAGATGGAAGAAGATTTTGAA
ATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAGCCGTTATTACTTTATTAA
AATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGATTTTGCACCAATTATTAA
GGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTATATCTATTTTAAGACACGT
TTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAGTCAGAAACGTCGTTTTAA
TAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTTTGTCAAGTGGCGGTATGC
AATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCATATAAACAAGGTCTGAGT
ACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGAGCACTTTTTGGGAAGGCA
AAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTAAGGAAAAATAGTTAAAAA
GGGAAAGAATGGAACATGTGATTGTACCATTCTTTTTGGTTGAAAATTAAGAAAAGTTA
TTATAAATTATTGGTTTAACATGCCATATTA

Figure 11

~~The comA and comB nucleotide and amino acid sequences. ComA and ComB are the components of the CSP exporter.~~

A.

~~S. mutans comA gene~~

ATGAAACAAGTTATTTATGTTGTTTTAATCGTCATAGCCGTTAACATTCTCTTAGAGAT
TATCAAAAGAGTAACAAAAAGGGGAGGGACAGTTTCGTCATCTAATCCTTTACCAGATG
GGCAGTCTAAGTTGTTTTGGCGCAGACATTATAAGCTAGTACCTCAGATTGATACCAGA
GACTGTGGGCCGGCAGTGCTGGCATCTGTTGCAAAGCATTACGGATCTAATTACTCTAT
CGCTTATCTGCGGGAACTCTCAAAGACTAACAAGCAGGGAACAACAGCTCTTGGCATTG
TTGAAGCTGCTAAAAAGTTAGGCTTTGAAACACGCTCTATCAAGGCGGATATGACGCTT
TTTGATTATAATGATTTGACCTATCCTTTTATCGTCCATGTGATTAAAGGAAAACGTCT
GCAGCATTATTATGTCGTCTATGGCAGCCAGAATAATCAGCTGATTATTGGAGATCCTG
ATCCTTCAGTTAAGGTGACTAGGATGAGTAAGGAACGCTTTCAATCAGAGTGGACAGGC
CTTGCAATTTTCCTAGCTCCTCAGCCTAACTATAAGCCTCATAAAGGTGAAAAAAATGG
TTTGTCTAATTTCTTCCCGTTGATCTTTAAGCAGAAAGCTTTGATGACTTATATTATCA
TAGCTAGCTTGATTGTGACGCTCATTGATATTGTCGGATCATACTATCTCCAAGGAATA
TTGGACGAGTACATTCCTGATCAGCTGATTTCAACTTTAGGAATGATTACGATTGGTCT
GATAATAACCTATATTATCCAGCAGGTCATGGCTTTTGCAAAAGAATACCTCTTGGCCG
TACTCAGTTTGCGTTTAGTCATTGATGTTATCCTGTCTTATATCAAACATATTTTTACG
CTTCCTATGTCTTTCTTTGCGACAAGGCGAACAGGAGAAATCACGTCTCGTTTTACAGA
TGCCAATCAGATTATTGATGCTGTAGCGTCAACCATCTTTTCAATCTTTTTAGATATGA
CTATGGTAATTTTGGTTGGTGGGGTTTTGTTGGCGCAAAACAATAACCTTTTCTTTCTA
ACCTTGCTCTCCATTCCGATTTATGCCATCATTATTTTTGCTTTCTTGAAACCCTTTGA
GAAAATGAATCACGAAGTGATGGAAAGCAATGCTGTGGTAAGTTCTTCTATCATTGAAG
ATATCAATGGGATGGAAACCATTAAATCACTCACAAGTGAGTCCGCTCGTTATCAAAAC
ATTGATAGTGAATTTGTTGATTATTTGGAGAAAAACTTTAAGCTACACAAGTATAGTGC
CATTCAAACCGCATTAAAAAGCGGTGCTAAGCTTATCCTCAATGTTGTCATTCTCTGGT
ATGGCTCTCGTCTAGTTATGGATAATAAAATCTCAGTTGGTCAGCTTATCACCTTTAAT
GCTTTGCTGTCTTATTTCTCAAATCCAATTGAAAATATTATCAATCTGCAATCCAAACT
GCAGTCAGCTCGCGTTGCCAATACACGTCTTAATGAGGTCTATCTTGTCGAATCTGAAT
TTGAAAAAGACGGCGATTTATCAGAAAATAGCTTTTTAGATGGTGATATTTCGTTTGAA
AATCTTTCTTATAAATATGGATTTGGGCGAGATACCTTATCAGATATTAATTTATCAAT
CAAAAAAGGCTCCAAGGTCAGTCTAGTTGGAGCCAGTGGTTCTGGTAAAACAACTTTGG
CTAAACTGATTGTCAATTTCTACGAGCCTAACAAGGGGATTGTTCGAATCAATGGCAAT
GATTTAAAAGTTATTGATAAGACAGCTTTGCGGCGGCATATTAGCTATTTGCCGCAACA
GGCCTATGTTTTTAGTGGCTCTATTATGGATAATCTCGTTTTAGGAGCTAAAGAAGGAA
CGAGTCAGGAAGACATTATTCGTGCTTGTGAAATTGCTGAAATCCGCTCGGACATTGAA
CAAATGCCTCAGGGCTATCAGACAGAGTTATCAGATGGTGCCGGTATTTCTGGCGGTCA
AAAACAGCGGATTGCTTTAGCTAGGGCCTTATTAACACAGGCACCGGTTTTGATTCTGG
ATGAAGCCACCAGCAGTCTTGATATTTTGACAGAAAAGAAAATTATCAGCAATCTCTTA
CAGATGACGGAGAAAACAATAATTTTTGTTGCCCACCGCTTAAGCATTTCACAGCGTAC
TGACGAAGTCATTGTCATGGATCAGGGAAAAATTGTTGAACAAGGCACTCATAAGGAAC
TTTTAGCTAAGCAAGGTTTCTATTATAACCTGTTTAAT

B.

~~S. mutans ComA protein~~

Figure 11 (cont'd)

(B.)

MKQVIYVVLIVIAVNILLEIIKRVTKRGGTVSSSNPLPDGQSKLFWRRHYKLVPQIDTR
DCGPAVLASVAKHYGSNYSIAYLRELSKTNKQGTTALGIVEAAKKLGFETRSIKADMTL
FDYNDLTYPFIVHVIKGKRLQHYYVVYGSQNNQLIIGDPDPSVKVTRMSKERFQSEWTG
LAIFLAPQPNYKPHKGEKNGLSNFFPLIFKQKALMTYIIIASLIVTLIDIVGSYYLQGI
LDEYIPDQLISTLGMITIGLIITYIIQQVMAFAKEYLLAVLSLRLVIDVILSYIKHIFT
LPMSFFATRRTGEITSRFTDANQIIDAVASTIFSIFLDMTMVILVGGVLLAQNNNLFFL
TLLSIPIYAIIIFAFLKPFEKMNHEVMESNAVVSSSIIEDINGMETIKSLTSESARYQN
IDSEFVDYLEKNFKLHKYSAIQTALKSGAKLILNVVILWYGSRLVMDNKISVGQLITFN
ALLSYFSNPIENIINLQSKLQSARVANTRLNEVYLVESEFEKDGDLSENSFLDGDISFE
NLSYKYGFGRDTLSDINLSIKKGSKVSLVGASGSGKTTLAKLIVNFYEPNKGIVRINGN
DLKVIDKTALRRHISYLPQQAYVFSGSIMDNLVLGAKEGTSQEDIIRACEIAEIRSDIE
QMPQGYQTELSDGAGISGGQKQRIALARALLTQAPVLILDEATSSLDILTEKKIISNLL
QMTEKTIIFVAHRLSISQRTDEVIVMDQGKIVEQGTHKELLAKQGFYYNLFN

(C.)

~~S. mutans comB gene~~

ATGGATCCTAAATTTTTACAAAGTGCAGAATTTTATAGGAGACGCTATCATAATTTTGC
GACACTATTAATTGTTCCTTTGGTCTGCTTGATTATCTTCTTGGTCATATTCCTTTGTT
TTGCTAAAAAAGAAATTACAGTGATTTCTACTGGTGAAGTTGCACCAACAAAGGTTGTA
GATGTTATCCAATCTTACAGTGACAGTTCAATCATTAAAAATAATTTAGATAATAATGC
AGCTGTTGAGAAGGGAGACGTTTTAATTGAATATTCAGAAAATGCCAGTCCAAACCGTC
AGACTGAACAAAAGAATATTATAAAAGAAAGACAAAAACGAGAAGAGAAGGAAAAGAAA
AAACACCAAAAGAGCAAGAAAAAGAAGAAGTCTAAGAGCAAGAAAGCTTCCAAAGATAA
GAAAAAGAAATCGAAAGACAAGGAAAGCAGCTCTGACGATGAAAATGAGACAAAAAAGG
TTTCGATTTTTGCTTCAGAAGATGGTATTATTCATACCAATCCCAAATATGATGGTGCC
AATATTATTCCGAAGCAAACCGAGATTGCTCAAATCTATCCTGATATTCAAAAAACAAG
AAAAGTGTTAATCACCTATTATGCTTCTTCTGATGATGTTGTTTCTATGAAAAAGGGGC
AAACCGCTCGTCTTTCCTTGGAAAAAAAGGGAAATGACAAGGTTGTTATTGAAGGAAAA
ATTAACAATGTCGCTTCATCAGCAACTACTACTAAAAAAGGAAATCTCTTTAAGGTTAC
TGCCAAAGTAAAGGTTTCTAAGAAAAATAGCAAACTCATCAAGTATGGTATGACAGGCA
AGACAGTCACTGTCATTGATAAAAAGACTTATTTTGATTATTTCAAAGATAAATTACTG
CATAAAATGGATAAT

(D.)

~~S. mutans ComB protein~~

MDPKFLQSAEFYRRRYHNFATLLIVPLVCLIIFLVIFLCFAKKEITVISTGEVAPTKVV
DVIQSYSDSSIIKNNLDNNAAVEKGDVLIEYSENASPNRQTEQKNIIKERQKREEKEKK
KHQKSKKKKKSKSKKASKDKKKKSKDKESSSDDENETKKVSIFASEDGIIHTNPKYDGA
NIIPKQTEIAQIYPDIQKTRKVLITYYASSDDVVSMKKGQTARLSLEKKGNDKVVIEGK
INNVASSATTTKKGNLFKVTAKVKVSKKNSKLIKYGMTGKTVTVIDKKTYFDYFKDKLL
HKMDN

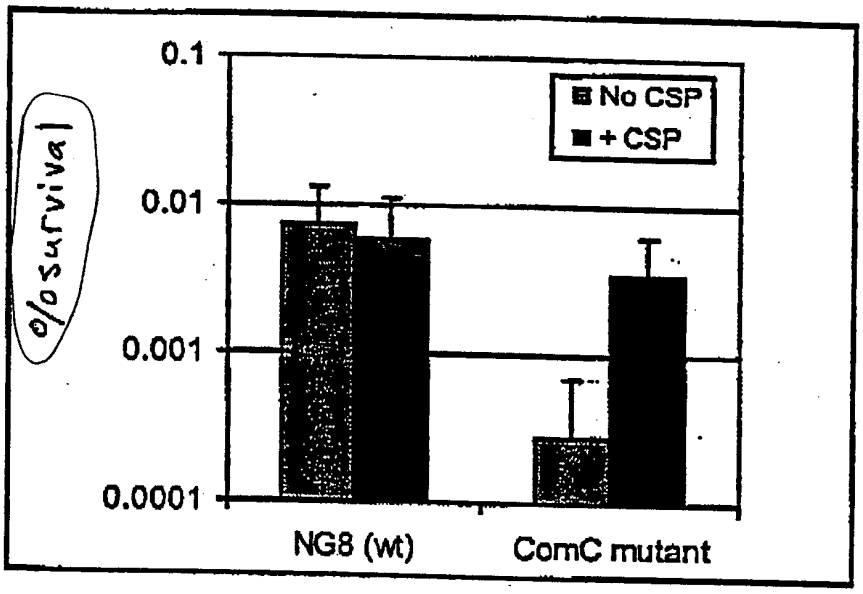

Figure 12